

# SEARCH REQUEST FORM

Requestor's

Name: \_\_\_\_\_

Serial

Number: \_\_\_\_\_

Date: \_\_\_\_\_

Phone: \_\_\_\_\_

Art Unit: \_\_\_\_\_

## Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

## STAFF USE ONLY

Date completed: 02-06-03

Search Site: \_\_\_\_\_

Vendors

Searcher: Beverly C 4994

STIC

IG Suite

Terminal time: 20

CM-1

STN

Elapsed time: \_\_\_\_\_

Pre-S

Dialog

CPU time: \_\_\_\_\_

Type of Search

APS

Total time: 25

N.A. Sequence

Geninfo

Number of Searches: \_\_\_\_\_

A.A. Sequence

SDC

Number of Databases: 1

Structure

DARC/Questel

Bibliographic

Other CGN

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 14:17:31 ; Search time 0.90249 Seconds  
(without alignments)  
1597.820 Million cell updates/sec

Title: US-09-689-159a-2\_COPY\_346\_360

Perfect score: 15

Sequence: 1 SHLGPHTSPESRAA 15

Scoring table: OLIGO

Searched: 283224 seqs, 96134422 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database: PIR\_73:\*

1: pir1:\*\n2: pir2:\*\n3: pir3:\*\n4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	100.0	463	2	presenilin 1-463 -
2	15	100.0	467	2	presenilin 1, spli
3	15	100.0	467	2	S182 protein - mou
4	8	53.3	347	2	hypothetical prote
5	7	46.7	327	2	hypothetical prote
6	7	46.7	343	2	hypothetical prote
7	7	46.7	463	2	probable transcrip
8	7	46.7	467	2	presenilin 1 prote
9	7	46.7	512	2	methionine-tRNA II
10	6	40.0	87	2	probable regulator
11	6	40.0	87	2	probable regulator
12	6	40.0	138	2	hypothetical prote
13	6	40.0	139	2	hypothetical prote
14	6	40.0	166	2	hypothetical prote
15	6	40.0	240	2	molybdenum cofacto
16	6	40.0	241	2	probable regulator
17	6	40.0	257	2	probable exodeoxyr
18	6	40.0	262	2	exodeoxyribonuclea
19	6	40.0	301	2	hypothetical prote
20	6	40.0	364	2	DNA replication an
21	6	40.0	382	2	opsin rh3 - fruit
22	6	40.0	401	2	multidrug-efflux t
23	6	40.0	407	2	conserved hypothet
24	6	40.0	408	2	histidyl-tRNA synt
25	6	40.0	429	2	histidine-tRNA lig
26	6	40.0	433	2	presenilin-alpha -
27	6	40.0	475	2	drop9 protein - hu
28	6	40.0	617	2	hypothetical prote
29	6	40.0	617	2	hypothetical prote

30	6	40.0	640	2	AE1895	hypothetical prote
31	6	40.0	806	2	I64207	preprolactin translo
32	6	40.0	817	1	RRVGCT	RNA-directed RNA p
33	6	40.0	818	1	RRVCCR	RNA-directed RNA p
34	6	40.0	851	2	S52717	RNA-directed RNA p
35	6	40.0	865	1	VGBERB	glycoprotein B pre
36	6	40.0	889	2	T30715	probable-major cor
37	6	40.0	981	2	B88794	protein K07f5.12a
38	6	40.0	1048	2	T30815	platelet-derived g
39	6	40.0	1098	1	PFMSRB	platelet-derived g
40	6	40.0	1106	1	PFHUGB	platelet-derived g
41	6	40.0	1680	2	T01367	hypothetical prote
42	6	40.0	1711	1	A47392	hypothetical prote
43	6	40.0	2664	2	T28626	chromodomain-helic
44	6	33.3	18	2	S58277	insulin-specific s
45	5	33.3	26	2	T04371	insulin-like growt
46	5	33.3	39	2	I17554	thymatin-like pro
47	5	33.3	51	2	F86491	homeobox - human
48	5	33.3	51	2	C72131	hypothetical prote
49	5	33.3	56	2	G72355	hypothetical prote
50	5	33.3	69	1	O0ECP7	ybea protein - Esc
51	5	33.3	69	2	C80713	hypothetical prote
52	5	33.3	69	2	G85563	hypothetical prote
53	5	33.3	76	2	D72667	hypothetical prote
54	5	33.3	82	2	S28280	hypothetical prote
55	5	33.3	87	2	AD3436	cytochrome c (vali
56	5	33.3	102	1	CCEG	conserved hypotnet
57	5	33.3	105	2	AB0318	Kora protein - Amy
58	5	33.3	118	2	S70089	hypothetical prote
59	5	33.3	122	2	A05114	MDH2 dehydrogenas
60	5	33.3	124	2	S43840	MDH2 dehydrogenas
61	6	33.3	127	2	AH3491	hypothetical cytos
62	6	33.3	134	2	S23003	trak protein - Esc
63	6	33.3	135	2	C62821	MDH2 dehydrogenas
64	5	33.3	136	2	B84990	50S ribosomal prot
65	5	33.3	145	2	AD2345	hypothetical prote
66	5	33.3	150	2	T08734	hypothetical prote
67	5	33.3	157	2	A97573	nitrogen regulator
68	5	33.3	160	2	A71062	hypothetical prote
69	5	33.3	163	2	B64172	hypothetical prote
70	5	33.3	165	2	D87241	glucokinase - rats
71	5	33.3	165	2	T16984	transcription fact
72	5	33.3	166	2	I49694	glucokinase - rats
73	5	33.3	166	2	I84740	glucokinase - rats
74	5	33.3	168	2	S68480	succinate dehydrog
75	5	33.3	173	2	F75274	conserved hypotnet
76	5	33.3	177	2	S14747	splicingmyelin phos
77	5	33.3	176	2	T36394	probable pantocate-
78	5	33.3	183	2	A11820	hypothetical prote
79	5	33.3	185	2	T01887	hypothetical prote
80	5	33.3	189	2	G90855	hypothetical prote
81	5	33.3	189	2	A86369	hypothetical prote
82	5	33.3	191	2	A82613	conserved hypotnet
83	5	33.3	191	2	T28682	hypothetical prote
84	5	33.3	197	2	B72471	hypothetical prote
85	5	33.3	198	2	A45067	laminin B1 chain v
86	5	33.3	201	2	A12631	hypothetical prote
87	5	33.3	201	2	B86488	hypothetical prote
88	5	33.3	206	2	T07771	probable heme tran
89	5	33.3	206	2	T09503	probable heme tran
90	5	33.3	206	2	AF0498	hypothetical prote
91	5	33.3	207	2	T17623	hypothetical prote
92	5	33.3	208	2	A48567	calmodulin-ubiquit
93	5	33.3	214	2	A82157	hypothetical prote
94	5	33.3	214	2	F82627	phosphoglycerate m
95	5	33.3	215	2	T16393	hypothetical prote
96	5	33.3	218	2	H83475	probable transcrip
97	5	33.3	219	2	T42605	envelope protein -
98	5	33.3	224	2	A87094	ATP-dependent clip
99	5	33.3	227	2	D85018	probable hypotnet
100	5	33.3	228	2	D95865	probable pentose-5
101	5	33.3	230	2	S41043	rum1 protein - fis
102	5	33.3	230	2	T40233	Rum1p - fission ye

103	5	33.3	230	2	A54501	sporulated coocyst
104	5	33.3	238	2	C72576	probable glutamine
105	5	33.3	244	2	B88115	protein F53c3.4 [i
106	5	33.3	244	2	AG2071	hypothetical prote
107	5	33.3	247	2	T33469	hypothetical prote
108	5	33.3	247	2	H69030	coenzyme PQQ synth
109	5	33.3	247	2	T17311	hypothetical prote
110	5	33.3	247	2	B97669	hypothetical prote
111	5	33.3	247	2	AG2893	conserved hypotet
112	5	33.3	249	2	A70745	probable transcrip
113	5	33.3	249	2	T16924	hypothetical prote
114	5	33.3	252	2	B48725	MOV specific prote
115	5	33.3	263	3	AG2910	hypothetical prote
116	5	33.3	265	2	A87345	transcription regu
117	5	33.3	266	2	A12762	conserved hypotet
118	5	33.3	267	2	B97685	monofunctional bio
119	5	33.3	268	2	AD0898	PTS-transport fam1
120	5	33.3	270	2	G75411	probable transposa
121	5	33.3	271	2	C96010	conserved hypotet
122	5	33.3	271	2	H97543	hypothetical prote
123	5	33.3	272	2	AG2215	hypothetical prote
124	5	33.3	274	2	T22993	hypothetical prote
125	5	33.3	275	1	G69963	lipoprotein SpoIII
126	5	33.3	278	1	B69025	conserved hypotet
127	5	33.3	279	2	B83456	hypothetical prote
128	5	33.3	280	2	C70642	probable ribosomal
129	5	33.3	281	2	T28857	hypothetical prote
130	5	33.3	284	2	T13621	hypothetical prote
131	5	33.3	286	2	B84807	probable RNA-bindi
132	5	33.3	287	2	T15779	hypothetical prote
133	5	33.3	291	2	G70605	hypothetical prote
134	5	33.3	292	2	T00996	En/Spm-like transp
135	5	33.3	295	2	A43663	host-inducible pro
136	5	33.3	295	2	AF0098	probable exported
137	5	33.3	297	2	T13317	hypothetical prote
138	5	33.3	299	2	F90741	hypothetical prote
139	5	33.3	299	2	A85592	hypothetical prote
140	5	33.3	299	2	AH0602	probable formate a
141	5	33.3	301	1	A54687	transcription fact
142	5	33.3	301	1	A48880	transcription fact
143	5	33.3	301	2	B84282	8-oxoguanine DNA g
144	5	33.3	302	2	T13457	hypothetical prote
145	5	33.3	302	2	F84329	hypothetical prote
146	5	33.3	302	2	S75481	polysialic acid tr
147	5	33.3	303	2	G64405	tetrahydromethanop
148	5	33.3	306	2	B70835	hypothetical prote
149	5	33.3	308	1	H64839	formate acetyltran
150	5	33.3	309	2	T41889	PE38 ori153 - Bomb
151	5	33.3	314	2	T26531	hypothetical prote
152	5	33.3	316	2	F87260	MeCB/TaqB/Cpsf fam
153	5	33.3	317	2	F82672	ATP sulfurylase, s
154	5	33.3	320	2	AB0330	phospholipase A [1
155	5	33.3	321	1	A43681	immediate-early pr
156	5	33.3	321	2	C72869	hypothetical prote
157	5	33.3	321	2	S55640	hypothetical prote
158	5	33.3	323	2	T13005	hypothetical prote
159	5	33.3	326	2	F96522	hypothetical prote
160	5	33.3	328	2	A84291	ornithine cyclodea
161	5	33.3	330	2	A40855	homeotic protein H
162	5	33.3	332	2	S44743	CO2D5.1 protein -
163	5	33.3	332	2	A70388	hydrogenase expres
164	5	33.3	332	2	I57032	gene Tlx-1 protein
165	5	33.3	332	2	AG0809	probable membrane
166	5	33.3	334	2	C22735	hypothetical nox2
167	5	33.3	334	2	A83342	conserved hypotet
168	5	33.3	335	2	B86429	F26G16.5 protein -
169	5	33.3	336	2	T44968	oxidoreductase [im
170	5	33.3	338	2	F83059	ketol-acid reducto
171	5	33.3	338	2	A83835	NADH oxidase BHL48
172	5	33.3	338	2	C98242	hypothetical prote
173	5	33.3	338	2	A13043	oxidoreductase Atu
174	5	33.3	342	2	B85757	partial probable o
175	5	33.3	342	2	AB3116	transcription regu

176	5	33.3	346	2	C75139	bifunctional short
177	5	33.3	346	2	T47411	hypothetical prote
178	5	33.3	348	2	G85638	hypothetical prote
179	5	33.3	349	2	T24015	hypothetical prote
180	5	33.3	351	2	D96761	unknown protein [1
181	5	33.3	353	2	F83190	probable UDP-3-O-1
182	5	33.3	356	2	H70882	hypothetical prote
183	5	33.3	356	2	C98171	hypothetical prote
184	5	33.3	358	2	T38920	prostaglandin E2 r
185	5	33.3	358	2	S51312	EP2 prostaglandin
186	5	33.3	361	2	G87706	hypothetical prote
187	5	33.3	363	2	S33702	homeotic protein D
188	5	33.3	364	2	AD3557	acrilflavin resista
189	5	33.3	365	2	A34424	CD4 membrane glyco
190	5	33.3	366	2	S51363	actin modulator pr
191	5	33.3	368	2	G83180	probable FMN oxido
192	5	33.3	371	2	AB0407	probable zinc-bind
193	5	33.3	371	2	T27643	hypothetical prote
194	5	33.3	374	2	AD0719	probable bacteriop
195	5	33.3	375	2	T23705	hypothetical prote
196	5	33.3	377	2	S21302	succinate dehydrog
197	5	33.3	377	2	H71081	hypothetical prote
198	5	33.3	377	2	UC7535	chitinase (EC 3.2.
199	5	33.3	379	2	S46711	hypothetical prote
200	5	33.3	380	1	TVMSF	transforming prote
201	5	33.3	380	2	T04284	hypothetical prote
202	5	33.3	381	1	TVMW7	transforming prote
203	5	33.3	384	1	T22434	phosphoprotein pho
204	5	33.3	384	2	AF0636	glucans biosynthes
205	5	33.3	387	2	S65959	cytochrome-c perox
206	5	33.3	388	1	WMVZM1	major envelope ant
207	5	33.3	388	1	WMVZU2	major envelope ant
208	5	33.3	388	2	T30623	major envelope ant
209	5	33.3	388	2	G95375	probable drug resi
210	5	33.3	394	2	A55045	probable 3-hydroxy
211	5	33.3	398	1	S47520	vitamin D-3 25-hyd
212	5	33.3	399	2	F87356	hypothetical prote
213	5	33.3	401	2	T36661	probable DNA polym
214	5	33.3	403	2	B35401	cytochrome P450 10
215	5	33.3	403	2	B70961	probable esterase
216	5	33.3	403	2	T39697	DNAJ protein - fls
217	5	33.3	404	2	S75529	beta ketoacyl-acyl
218	5	33.3	405	2	C96640	hypothetical prote
219	5	33.3	405	4	A61181	homeotic protein H
220	5	33.3	411	2	C87586	metal ion efflux m
221	5	33.3	413	2	A34888	transcription fact
222	5	33.3	413	2	S41114	Gal beta 1,4 GlcNA
223	5	33.3	414	2	A12544	hypothetical prote
224	5	33.3	416	2	E82231	Nifs-related prote
225	5	33.3	417	2	T34561	hypothetical prote
226	5	33.3	418	2	T36358	hypothetical prote
227	5	33.3	422	2	T33562	hypothetical prote
228	5	33.3	425	1	JH0710	transcription fact
229	5	33.3	425	2	AF2246	phosphoribosylform
230	5	33.3	430	1	UYPVE1	nonstructural prot
231	5	33.3	430	2	T21060	hypothetical prote
232	5	33.3	434	2	S50865	avermectin-sensitl
233	5	33.3	435	2	A46231	helix-loop-helix p
234	5	33.3	437	2	T42653	hypothetical prote
235	5	33.3	438	2	A46259	recombination prot
236	5	33.3	439	2	T01050	hypothetical prote
237	5	33.3	439	2	E87638	omega-amino acid-p
238	5	33.3	439	2	F82997	two-component sens
239	5	33.3	442	2	T30157	hypothetical prote
240	5	33.3	442	2	T20788	hypothetical prote
241	5	33.3	445	2	B82954	probable dihydooor
242	5	33.3	447	2	T50705	gamma-aminobutyrat
243	5	33.3	448	2	B85077	hypothetical prote
244	5	33.3	451	2	T36217	plasmid transfer p
245	5	33.3	451	2	G69070	coenzyme F390 synt
246	5	33.3	454	2	T18765	hypothetical prote
247	5	33.3	457	2	S44269	platelet-derived g
248	5	33.3	457	2	A38567	GTP-binding protei



249	5	33.3	457	2	A99862	probable outer mem	322	5	33.3	596	2	JC4697	trehalose trehaloh
250	5	33.3	459	2	AC0075	probable membrane	323	5	33.3	599	2	S75048	aspartate-tRNA lig
251	5	33.3	459	2	T39473	probable geranylge	324	5	33.3	601	2	C89451	protein T04G.6 [l
252	5	33.3	462	2	T11637	hypothetical prote	325	5	33.3	602	2	G81195	asparlyl-tRNA synt
253	5	33.3	464	2	C46157	hexokinase (EC 2.7	326	5	33.3	602	2	F81831	aspartate-tRNA lig
254	5	33.3	465	2	A46157	hexokinase (EC 2.7	327	5	33.3	604	2	T24456	hypothetical prote
255	5	33.3	465	2	A31810	hexokinase (EC 2.7	328	5	33.3	605	2	T43974	hypothetical prote
256	5	33.3	465	2	I49693	glucokinase (EC 2.	329	5	33.3	610	2	S76234	hypothetical prote
257	5	33.3	466	2	D46157	hexokinase (EC 2.7	330	5	33.3	610	2	S50539	hypothetical prote
258	5	33.3	467	2	D86672	efflux pump antidi	331	5	33.3	611	1	S12566	translation initia
259	5	33.3	468	2	C82722	UDP-N-acetylmuramo	332	5	33.3	611	2	S52229	hypothetical prote
260	5	33.3	471	1	SYECET	glutamate-tRNA lig	333	5	33.3	612	2	AC2396	hypothetical prote
261	5	33.3	471	1	F91038	glutamate-tRNA syn	334	5	33.3	613	2	T00077	gag-like protein -
262	5	33.3	471	2	H85882	glutamate tRNA syn	335	5	33.3	613	2	T21325	hypothetical prote
263	5	33.3	471	2	T46860	probable glutamate	336	5	33.3	613	2	T34011	hypothetical prote
264	5	33.3	471	2	B83127	probable secretion	337	5	33.3	617	2	T51708	methylyltransferase
265	5	33.3	471	2	S46739	hypothetical prote	338	5	33.3	627	2	D84531	probable salt-indu
266	5	33.3	474	2	A10808	glutamate-tRNA lig	339	5	33.3	628	2	S01955	hypothetical prote
267	5	33.3	476	2	T43464	hypothetical prote	340	5	33.3	628	2	S19150	hypothetical prote
268	5	33.3	480	2	B64059	glutamate-tRNA lig	341	5	33.3	629	2	F84428	probable myosin he
269	5	33.3	483	2	T16443	hypothetical prote	342	5	33.3	629	2	T48940	hypothetical prote
270	5	33.3	484	1	SYRZET	glutamate-tRNA lig	343	5	33.3	630	2	F63073	hypothetical prote
271	5	33.3	484	1	T36427	probable rhamnose	344	5	33.3	632	2	B69310	hypothetical prote
272	5	33.3	484	2	T50909	hypothetical prote	345	5	33.3	634	2	T49415	hypothetical prote
273	5	33.3	484	2	D84725	probable acetylact	346	5	33.3	638	2	S67605	hypothetical prote
274	5	33.3	497	2	T51195	hypothetical prote	347	5	33.3	638	2	JC7753	ring finger B-box
275	5	33.3	498	1	A56564	transcription fact	348	5	33.3	641	2	A45054	probable intercell
276	5	33.3	502	2	T02306	probable protein k	349	5	33.3	644	2	T10403	p74 protein - Orgy
277	5	33.3	505	2	S15194	sucrose porin scry	350	5	33.3	645	2	H96011	asparagine synthas
278	5	33.3	506	2	E82721	lysyl-tRNA synthet	351	5	33.3	647	2	C69102	DNA mismatch recog
279	5	33.3	506	2	T41623	probable protein k	352	5	33.3	651	2	B41288	vascular cell adhe
280	5	33.3	508	2	A70125	UDP-N-acetylmuramo	353	5	33.3	651	2	D86977	probable primosoma
281	5	33.3	508	2	AD0076	probable membrane	354	5	33.3	658	2	C81860	DNA mismatch repai
282	5	33.3	509	2	S74935	hypothetical prote	355	5	33.3	662	2	B81084	mismatch repair pr
283	5	33.3	512	2	H84320	phycoene dehydroge	356	5	33.3	662	2	G87630	hypothetical prote
284	5	33.3	513	2	T38044	hypothetical prote	357	5	33.3	665	2	T13744	hypothetical prote
285	5	33.3	515	2	S67290	probable membrane	358	5	33.3	668	2	JC2363	protein kinase (EC
286	5	33.3	515	2	T43152	hypothetical prote	359	5	33.3	669	2	F84377	mismatch repair pr
287	5	33.3	516	2	S28060	serum response fac	360	5	33.3	671	2	D86479	hypothetical prote
288	5	33.3	517	2	B87644	4-coumarate-CoA li	361	5	33.3	673	1	AQH068	annexin VI [valida
289	5	33.3	521	2	T06606	hypothetical prote	362	5	33.3	673	1	S01786	annexin VI - rat
290	5	33.3	530	2	A84922	hypothetical prote	363	5	33.3	673	1	S52844	TPR domain protein
291	5	33.3	539	2	T30327	hsdm protein - Kie	364	5	33.3	673	2	F87636	nuclear autoantige
292	5	33.3	543	2	B34612	zinc finger protei	365	5	33.3	680	2	A43800	rabphillin-3A - mou
293	5	33.3	544	2	Aei563	pyruvate dehydroge	366	5	33.3	681	2	JX0338	rabphillin-3A - rat
294	5	33.3	546	2	AP1206	pyruvate dehydroge	367	5	33.3	684	2	I58166	nuclear protein NO
295	5	33.3	546	2	T24679	hypothetical prote	368	5	33.3	685	2	S46365	hypothetical prote
296	5	33.3	547	2	T36550	mercury(II) reduct	369	5	33.3	694	2	C72761	hypothetical prote
297	5	33.3	548	1	S70146	probable amino aci	370	5	33.3	696	2	T46394	Hsp83 protein - Le
298	5	33.3	549	2	D95979	hypothetical prote	371	5	33.3	700	2	S57415	hypothetical prote
299	5	33.3	550	2	T19257	iron transport pro	372	5	33.3	701	2	F84912	hypothetical prote
300	5	33.3	552	2	S50313	hypothetical prote	373	5	33.3	704	2	T24517	probable DNA helic
301	5	33.3	559	2	F86479	nuclear receptor R	374	5	33.3	721	2	S57374	probable beta-gala
302	5	33.3	560	2	JC2494	hypothetical prote	375	5	33.3	722	2	T45495	gene CARSR12 prote
303	5	33.3	560	2	C71155	hemagglutinin prec	376	5	33.3	722	2	T02421	probable beta-gala
304	5	33.3	562	1	HMIV2	SIS7 protein - yea	377	5	33.3	722	2	S27043	hypothetical prote
305	5	33.3	562	2	S38149	beta-D-glucosidase	378	5	33.3	727	2	I56506	neurotransmitter t
306	5	33.3	563	2	T02720	hypothetical prote	379	5	33.3	727	2	T47541	Na+/Cl(-)-dependen
307	5	33.3	566	2	T45626	hypothetical prote	380	5	33.3	727	2	T25362	beta-galactosidase
308	5	33.3	567	2	C84425	similar to axl 1 p	381	5	33.3	729	2	T04269	probable beta-gala
309	5	33.3	567	2	T15574	hypothetical prote	382	5	33.3	731	2	S16595	gene CARSR12 prote
310	5	33.3	568	2	C72129	probable outer mem	383	5	33.3	731	2	T17002	probable beta-gala
311	5	33.3	568	2	G86493	probable leader pe	384	5	33.3	732	2	T25937	hypothetical prote
312	5	33.3	569	2	T17779	hypothetical prote	385	5	33.3	736	2	A99279	hypothetical prote
313	5	33.3	570	2	T27407	membrane associate	386	5	33.3	736	2	T40080	probable ABC trans
314	5	33.3	570	2	H97244	aspartate-tRNA lig	387	5	33.3	736	2	T40510	beta transducin -
315	5	33.3	579	2	F72252	zinc finger protei	388	5	33.3	739	2	A41288	vascular cell adhe
316	5	33.3	579	2	T34331	probable serine/th	389	5	33.3	741	2	B54908	phospholipase A2 c
317	5	33.3	586	1	S47452	aspartate-tRNA lig	390	5	33.3	741	2	A26572	bsg25D protein - f
318	5	33.3	589	1	S42047	aspartate-tRNA lig	391	5	33.3	751	2	S38101	hypothetical prote
319	5	33.3	593	1	T37295	GTP-binding protei	392	5	33.3	753	2	C97510	hypothetical prote
320	5	33.3	595	2	AE2110	aspartate-tRNA lig	393	5	33.3	753	2	AB2729	succinylglycan bios
321	5	33.3	596	2	C70724	aspartate-tRNA lig	394	5	33.3	754	2	AH3004	vgrc protein [limpo

395	5	33.3	755	2	T40553	468	5	33.3	1061	2	D69799	cytochrome P450 /
396	5	33.3	756	2	G84866	469	5	33.3	1067	2	AC1439	B. subtilis yueB p
397	5	33.3	757	2	F75590	470	5	33.3	1071	2	D86279	hypothetical prote
398	5	33.3	760	2	J01383	471	5	33.3	1075	2	T07448	probable DNA-direc
399	5	33.3	765	2	AD0778	472	5	33.3	1088	1	PFRUGA	platelet-derived g
400	5	33.3	765	2	C64981	473	5	33.3	1089	1	PFRUGA	platelet-derived g
401	5	33.3	765	2	C91006	474	5	33.3	1089	1	S33727	platelet-derived g
402	5	33.3	765	2	D85806	475	5	33.3	1101	2	T33153	hypothetical prote
403	5	33.3	765	2	T35719	476	5	33.3	1104	2	S59310	probable membrane
404	5	33.3	784	2	T26585	477	5	33.3	1112	2	S46313	phytochrome E - Ar
405	5	33.3	786	2	T39585	478	5	33.3	1114	2	T14351	serine/threonine-s
406	5	33.3	787	1	JDYLM2	479	5	33.3	1121	2	I38127	phosphoprotein pho
407	5	33.3	788	1	JDVYLD	480	5	33.3	1121	2	T02764	myosin-I binding p
408	5	33.3	790	2	G02678	481	5	33.3	1127	2	T25804	hypothetical prote
409	5	33.3	792	2	G86564	482	5	33.3	1134	1	JN0711	protein-tyrosine k
410	5	33.3	792	2	D81608	483	5	33.3	1135	2	J02163	hypothetical 126.5
411	5	33.3	792	2	H72058	484	5	33.3	1138	1	S24066	protein-tyrosine k
412	5	33.3	799	2	T02456	485	5	33.3	1139	1	H64089	DNA-directed DNA p
413	5	33.3	808	2	S73947	486	5	33.3	1160	1	DJEC3A	DNA-directed DNA p
414	5	33.3	810	2	T33323	487	5	33.3	1160	1	A45915	DNA-directed DNA p
415	5	33.3	813	2	T02672	488	5	33.3	1160	2	B90652	DNA polymerase III
416	5	33.3	814	1	T39627	489	5	33.3	1160	2	B85503	DNA polymerase III
417	5	33.3	815	2	T57487	490	5	33.3	1160	2	A10530	DNA polymerase III
418	5	33.3	816	2	B98196	491	5	33.3	1164	2	G82100	DNA polymerase III
419	5	33.3	816	2	S16328	492	5	33.3	1171	2	AB0130	DNA-directed DNA p
420	5	33.3	816	2	AH3090	493	5	33.3	1173	2	I50620	nuclear envelope p
421	5	33.3	817	2	T25758	494	5	33.3	1199	2	A40670	prockr2 - chicken
422	5	33.3	824	2	T20477	495	5	33.3	1226	2	S15053	3c3.20c protein -
423	5	33.3	831	2	T05771	496	5	33.3	1321	2	T10929	dextranase - Strept
424	5	33.3	832	2	E84543	497	5	33.3	1337	2	T30291	LamDR1 protein - L
425	5	33.3	832	2	S41889	498	5	33.3	1341	2	JG0166	hypothetical prote
426	5	33.3	833	2	T24682	499	5	33.3	1341	2	T17285	hypothetical prote
427	5	33.3	835	2	T06590	500	5	33.3	1345	2	T00964	protein-tyrosine k
428	5	33.3	838	2	T40203	501	5	33.3	1404	1	A48196	hypothetical prote
429	5	33.3	839	2	C84685	502	5	33.3	1411	2	T18417	DNA polymerase III
430	5	33.3	842	2	H86220	503	5	33.3	1415	2	S52267	hypothetical prote
431	5	33.3	845	2	I48176	504	5	33.3	1417	2	T18418	DNA polymerase III
432	5	33.3	858	2	A46613	505	5	33.3	1438	2	C89900	DNA polymerase III
433	5	33.3	864	2	H85335	506	5	33.3	1474	2	B85188	retrotransposon II
434	5	33.3	864	2	T04518	507	5	33.3	1493	2	A38218	GAP-associated pro
435	5	33.3	875	2	AE1270	508	5	33.3	1499	2	B59431	Rho GTPase activat
436	5	33.3	875	2	AC1632	509	5	33.3	1501	2	T42724	p190-B protein - m
437	5	33.3	877	2	S49783	510	5	33.3	1535	2	T49042	hypothetical prote
438	5	33.3	880	2	S60137	511	5	33.3	1685	2	T43217	RNA polymerase (EC
439	5	33.3	886	2	C87031	512	5	33.3	1729	2	A49282	fusion protein Ia/
440	5	33.3	888	2	A55318	513	5	33.3	1751	2	G71518	hypothetical prote
441	5	33.3	888	2	JC5399	514	5	33.3	1771	2	S76851	hypothetical prote
442	5	33.3	913	2	T15278	515	5	33.3	1795	2	F97713	190K antigen precu
443	5	33.3	922	2	AG1827	516	5	33.3	1926	2	S01169	beta-glycosidase c
444	5	33.3	924	2	T09220	517	5	33.3	1994	2	D86452	protein F6N18.13 l
445	5	33.3	928	2	T10164	518	5	33.3	2088	2	E71436	hypothetical prote
446	5	33.3	931	2	D86222	519	5	33.3	2240	2	T37057	probable multi-dom
447	5	33.3	938	2	H83050	520	5	33.3	2288	2	T30568	acetyl-CoA carboxy
448	5	33.3	940	1	BVECUA	521	5	33.3	2332	1	GNNYF	genome polyprotein
449	5	33.3	940	2	H91258	522	5	33.3	2332	1	GNNY4F	genome polyprotein
450	5	33.3	940	2	D86099	523	5	33.3	2333	1	GNNY2F	genome polyprotein
451	5	33.3	955	2	A47334	524	5	33.3	2336	2	S37077	genome polyprotein
452	5	33.3	956	2	H81654	525	5	33.3	2342	2	T13412	hypothetical prote
453	5	33.3	957	2	T10633	526	5	33.3	2499	1	A30788	mannose 6-phosphat
454	5	33.3	965	2	S62935	527	5	33.3	2554	1	TVEF7L	kinase-related pro
455	5	33.3	968	2	T00353	528	5	33.3	2783	2	T34416	hypothetical prote
456	5	33.3	980	2	E72035	529	5	33.3	2793	2	B90784	hypothetical prote
457	5	33.3	980	2	E86589	530	5	33.3	2806	2	D85644	hypothetical prote
458	5	33.3	1003	2	T26746	531	5	33.3	3450	2	T26963	hypothetical prote
459	5	33.3	1008	2	T12532	532	5	33.3	3461	2	T26964	lysosomal traffick
460	5	33.3	1008	2	AE2304	533	5	33.3	3766	2	T18514	hypothetical prote
461	5	33.3	1021	2	A86421	534	5	33.3	3890	2	C89921	hypothetical prote
462	5	33.3	1021	2	T10748	535	5	33.3	4307	2	T20721	BIR repeat contain
463	5	33.3	1026	2	C71490	536	5	33.3	4845	10	T31067	Ig heavy chain CRD
464	5	33.3	1036	2	S55984	537	4	26.7	13	2	PT0309	T-cell antigen rec
465	5	33.3	1038	2	S06046	538	4	26.7	13	2	S47388	histone H2B - mous
466	5	33.3	1045	2	F90365	539	4	26.7	18	2	B27504	6-phosphofructo-2-
467	5	33.3	1054	2	D96519	540	4	26.7	20	2	I46940	

541	4	26.7	20	2	T44453	614	4	26.7	69	2	D70734	hypothetical prote
542	4	26.7	20	2	S78759	615	4	26.7	69	2	S41000	hypothetical prote
543	4	26.7	23	4	JH0515	616	4	26.7	70	2	S42895	env polyprotein -
544	4	26.7	26	4	I59183	617	4	26.7	70	2	JU0103	hypothetical 8.5K
545	4	26.7	27	2	A49236	618	4	26.7	70	2	B46522	T-cell receptor et
546	4	26.7	27	2	I50161	619	4	26.7	70	2	C46522	probable mbth prot
547	4	26.7	29	2	S17147	620	4	26.7	71	1	H70587	envelope glycoprot
548	4	26.7	29	2	S77569	621	4	26.7	71	2	S42909	env polyprotein -
549	4	26.7	33	2	A60601	622	4	26.7	71	2	S42906	env polyprotein -
550	4	26.7	35	2	PC3291	623	4	26.7	71	2	S42908	env polyprotein -
551	4	26.7	37	2	G49050	624	4	26.7	71	2	S42910	env polyprotein -
552	4	26.7	38	2	E82424	625	4	26.7	71	2	S42912	env polyprotein -
553	4	26.7	38	2	B31194	626	4	26.7	71	2	S42916	env polyprotein -
554	4	26.7	39	2	B36207	627	4	26.7	71	2	S42917	env polyprotein -
555	4	26.7	41	2	S09645	628	4	26.7	71	2	S42918	env polyprotein -
556	4	26.7	42	2	G86110	629	4	26.7	71	2	S42918	env polyprotein -
557	4	26.7	42	2	T01609	630	4	26.7	71	2	S42897	env polyprotein -
558	4	26.7	44	2	S29785	631	4	26.7	71	2	S42901	env polyprotein -
559	4	26.7	44	2	AB1046	632	4	26.7	71	2	S42903	env polyprotein -
560	4	26.7	45	2	T37124	633	4	26.7	71	2	S42904	env polyprotein -
561	4	26.7	48	2	S62189	634	4	26.7	71	2	S42905	env polyprotein -
562	4	26.7	49	2	B56448	635	4	26.7	71	2	T38721	env polyprotein -
563	4	26.7	51	2	T08488	636	4	26.7	71	2	E82610	HLA-DRB sigma anti
564	4	26.7	51	2	T29481	637	4	26.7	72	2	T30373	hypothetical prote
565	4	26.7	53	2	H82784	638	4	26.7	72	2	G33172	C-ORF-B protein -
566	4	26.7	54	2	I46202	639	4	26.7	72	2	A42856	EPF autoantibody-r
567	4	26.7	54	2	T35239	640	4	26.7	72	2	B64299	probable archaeal
568	4	26.7	56	2	A82565	641	4	26.7	72	2	I79659	DRB1 transplanta
569	4	26.7	57	2	D49056	642	4	26.7	73	2	D90864	phage shock protei
570	4	26.7	57	2	B95384	643	4	26.7	73	2	E85754	phage shock protei
571	4	26.7	58	2	A56448	644	4	26.7	73	2	S17124	probable ribosomal
572	4	26.7	58	2	T08023	645	4	26.7	73	2	T05930	hypothetical prote
573	4	26.7	58	2	AC1517	646	4	26.7	73	2	T17549	hypothetical prote
574	4	26.7	58	2	AP1158	647	4	26.7	74	2	H86993	hypothetical prote
575	4	26.7	59	2	T03588	648	4	26.7	74	2	AD3131	hypothetical prote
576	4	26.7	59	2	S21332	649	4	26.7	75	2	S35840	envelope protein -
577	4	26.7	59	2	T37115	650	4	26.7	75	2	S35806	envelope protein -
578	4	26.7	60	2	I55550	651	4	26.7	76	2	S35804	SNP2 protein - yea
579	4	26.7	60	2	C71245	652	4	26.7	76	2	T29087	hypothetical prote
580	4	26.7	60	2	E82516	653	4	26.7	77	2	T30471	hypothetical prote
581	4	26.7	60	2	AC2981	654	4	26.7	77	2	T17891	hypothetical prote
582	4	26.7	60	2	AE2514	655	4	26.7	77	2	A56194	chromoxane A-2 re
583	4	26.7	61	2	A36916	656	4	26.7	77	2	B87276	hypothetical prote
584	4	26.7	61	2	S18336	657	4	26.7	77	2	G83089	conserved hypotet
585	4	26.7	63	2	F82806	658	4	26.7	78	2	T24150	hypothetical prote
586	4	26.7	63	2	H82788	659	4	26.7	78	2	T29855	hypothetical prote
587	4	26.7	63	2	AG3289	660	4	26.7	79	2	F82602	conserved hypotet
588	4	26.7	64	2	JX0060	661	4	26.7	79	2	D43670	hypothetical prote
589	4	26.7	64	2	S11950	662	4	26.7	79	2	AD1825	hypothetical prote
590	4	26.7	64	2	T42161	663	4	26.7	80	2	AD2080	hypothetical prote
591	4	26.7	64	2	PN0648	664	4	26.7	80	2	S37112	class II histocomp
592	4	26.7	65	2	C87262	665	4	26.7	80	2	T34039	hypothetical prote
593	4	26.7	65	2	T17896	666	4	26.7	80	2	E71321	hypothetical prote
594	4	26.7	65	2	S75236	667	4	26.7	81	2	H86911	hypothetical prote
595	4	26.7	65	2	JN0802	668	4	26.7	81	2	S15142	hypothetical prote
596	4	26.7	65	2	AD3359	669	4	26.7	82	2	B83556	hypothetical prote
597	4	26.7	66	2	T36517	670	4	26.7	82	2	AI0032	probable type III
598	4	26.7	66	2	D72580	671	4	26.7	83	2	AC3940	hypothetical prote
599	4	26.7	66	2	AE1385	672	4	26.7	83	2	WM8PMB	gene p protein - p
600	4	26.7	66	2	AG1760	673	4	26.7	84	2	B95079	hypothetical prote
601	4	26.7	67	2	T10832	674	4	26.7	84	2	F81128	probable lipoprote
602	4	26.7	67	2	T16372	675	4	26.7	84	2	G97946	hypothetical prote
603	4	26.7	67	2	E96004	676	4	26.7	85	2	E64855	ycfr protein precu
604	4	26.7	67	2	H95418	677	4	26.7	85	2	F85674	hypothetical prote
605	4	26.7	67	2	A64321	678	4	26.7	85	2	B90815	hypothetical prote
606	4	26.7	67	2	D64416	679	4	26.7	85	2	A69985	hypothetical prote
607	4	26.7	67	2	D64513	680	4	26.7	85	2	G86484	10.0K hypothetical
608	4	26.7	68	2	JH0100	681	4	26.7	86	2	T42180	hypothetical prote
609	4	26.7	68	2	B96942	682	4	26.7				
610	4	26.7	68	2	AI1316	683	4	26.7				
611	4	26.7	69	2	H83236	684	4	26.7				
612	4	26.7	69	2	F83588	685	4	26.7				
613	4	26.7	69	2	J01703	686	4	26.7				

687	4	26.7	86	2	A57292	signal recognition
688	4	26.7	86	2	A34731	signal recognition
689	4	26.7	86	2	S57500	signal recognition
690	4	26.7	87	2	E82736	conserved hypotet
691	4	26.7	87	2	H84234	hypothetical prote
692	4	26.7	87	2	T30424	hypothetical prote
693	4	26.7	87	2	B97749	hypothetical prote
694	4	26.7	88	2	C84912	hypothetical prote
695	4	26.7	88	2	A84342	hypothetical prote
696	4	26.7	88	2	E97543	hypothetical prote
697	4	26.7	88	2	AE2762	conserved hypotet
698	4	26.7	89	1	C36869	probable export pr
699	4	26.7	89	2	S35277	probable export pr
700	4	26.7	89	2	S78699	probable export pr
701	4	26.7	89	2	AE0753	flagellar biosynth
702	4	26.7	89	2	A10221	flagellar biosynth
703	4	26.7	89	2	H90964	probable export pr
704	4	26.7	89	2	D87338	hypothetical prote
705	4	26.7	89	2	H85812	hypothetical prote
706	4	26.7	89	2	F97821	hypothetical prote
707	4	26.7	89	2	H86418	hypothetical prote
708	4	26.7	90	2	H82611	DNA-binding protei
709	4	26.7	91	2	J00903	bombayin B-1 homol
710	4	26.7	91	2	J00905	bombayin B-2 homol
711	4	26.7	91	2	I54248	apolipoprotein A-I
712	4	26.7	91	2	S02769	gag 75K protein pr
713	4	26.7	92	2	F83439	hypothetical prote
714	4	26.7	92	2	G69383	hypothetical prote
715	4	26.7	93	2	D48059	oncoprotein ZN-MYC
716	4	26.7	93	2	A13390	acyl carrier prote
717	4	26.7	94	2	S13265	heme oxygenase (de
718	4	26.7	94	2	C95249	PTS system, IIB co
719	4	26.7	94	2	F84497	En/Spm-like transp
720	4	26.7	94	2	T49851	hypothetical prote
721	4	26.7	94	2	H98113	hypothetical prote
722	4	26.7	95	2	AD2773	acgl carrier prote
723	4	26.7	95	2	A64422	calcium channel al
724	4	26.7	95	2	A70618	hypothetical prote
725	4	26.7	95	2	F84046	hypothetical prote
726	4	26.7	95	2	A97468	hypothetical prote
727	4	26.7	95	2	AD2686	hypothetical prote
728	4	26.7	96	2	T13540	hypothetical prote
729	4	26.7	97	2	AE2005	30S ribosomal prot
730	4	26.7	97	2	H84901	hypothetical prote
731	4	26.7	97	2	S43678	csnm protein - Chl
732	4	26.7	97	2	T28909	hypothetical prote
733	4	26.7	97	2	D64850	transcription fact
734	4	26.7	98	1	FERMX	ferredoxin [4Fe-4S
735	4	26.7	98	1	JC1001	ferredoxin [4Fe-4S
736	4	26.7	98	2	D95317	FlxX ferredoxin-11
737	4	26.7	98	2	I49562	alpha-1 type IIT c
738	4	26.7	98	2	E64763	probable alpha hel
739	4	26.7	98	2	D90680	probable alpha hel
740	4	26.7	98	2	E84589	hypothetical prote
741	4	26.7	98	2	A81902	hypothetical prote
742	4	26.7	98	2	B81120	hypothetical prote
743	4	26.7	98	2	S78727	protein YJL018c-a
744	4	26.7	98	2	H85530	probable alpha hel
745	4	26.7	99	1	WPBP2	repressor protein
746	4	26.7	99	2	S34301	galanin - mouse
747	4	26.7	99	2	C83116	50S ribosomal prot
748	4	26.7	99	2	S22957	hypothetical prote
749	4	26.7	99	2	B82720	hypothetical prote
750	4	26.7	100	2	A82333	primosomal replica
751	4	26.7	100	2	H70586	hypothetical prote
752	4	26.7	100	2	E36905	ribosomal protein
753	4	26.7	100	2	B72664	hypothetical prote
754	4	26.7	100	2	AE2335	hypothetical prote
755	4	26.7	101	2	A43262	hypothetical prote
756	4	26.7	101	2	S11936	saif protein - Stre
757	4	26.7	101	2	C90564	lipoprotein [impor
758	4	26.7	101	2	A72742	hypothetical prote
759	4	26.7	102	2	S15454	cytochrome c - Eug
760	4	26.7	102	2	C81682	conserved hypotet
761	4	26.7	102	2	C96034	conserved hypotet
762	4	26.7	102	2	A87422	hypothetical prote
763	4	26.7	102	2	B87669	hypothetical prote
764	4	26.7	102	2	G97493	hypothetical prote
765	4	26.7	102	2	A12711	hypothetical prote
766	4	26.7	103	2	A41403	hypothetical prote
767	4	26.7	103	2	S04101	aldosterone secret
768	4	26.7	103	2	T08501	hemolysin A - prot
769	4	26.7	103	2	D49852	trnd protein - Ent
770	4	26.7	103	2	F70976	hypothetical prote
771	4	26.7	103	2	F72690	hypothetical prote
772	4	26.7	103	2	A97484	hypothetical prote
773	4	26.7	103	2	A12701	hypothetical prote
774	4	26.7	103	2	S70182	hypothetical prote
775	4	26.7	103	2	S72973	hypothetical prote
776	4	26.7	104	1	JC4694	proteic killer act
777	4	26.7	104	1	S32568	NNO-12.3 protein -
778	4	26.7	104	2	E95299	hypothetical prote
779	4	26.7	104	2	C98209	hypothetical prote
780	4	26.7	104	2	AE3077	conserved hypotet
781	4	26.7	104	2	H91204	hypothetical prote
782	4	26.7	104	2	T09471	hypothetical prote
783	4	26.7	105	2	S46311	ferredoxin-nitrite
784	4	26.7	105	2	AE0581	conserved hypotet
785	4	26.7	105	2	B75392	hypothetical prote
786	4	26.7	105	2	C87324	hypothetical prote
787	4	26.7	105	2	S77115	hypothetical prote
788	4	26.7	106	2	AC0430	primosomal replica
789	4	26.7	106	2	A85078	probable DNA-direc
790	4	26.7	106	2	C86142	hypothetical prote
791	4	26.7	106	2	E75371	hypothetical prote
792	4	26.7	106	2	E70897	hypothetical prote
793	4	26.7	106	2	G84541	hypothetical prote
794	4	26.7	106	2	B64350	hypothetical prote
795	4	26.7	107	2	AH2877	hypothetical prote
796	4	26.7	107	2	D70626	hypothetical prote
797	4	26.7	107	2	A72701	hypothetical prote
798	4	26.7	108	2	C83205	conserved hypotet
799	4	26.7	108	2	C72587	hypothetical prote
800	4	26.7	108	2	C83528	hypothetical prote
801	4	26.7	108	2	T27503	hypothetical prote
802	4	26.7	109	1	CABO	hypothetical prote
803	4	26.7	109	2	JC2440	cytochrome-c oxida
804	4	26.7	109	2	T32588	phenyl dihydrogena
805	4	26.7	109	2	F72743	probable Arpase in
806	4	26.7	109	2	T13121	hypothetical prote
807	4	26.7	109	2	E70956	protein gp35 - pha
808	4	26.7	109	2	D75538	hypothetical prote
809	4	26.7	109	2	F72684	conserved hypotet
810	4	26.7	109	2	D72604	hypothetical prote
811	4	26.7	109	2	A87425	hypothetical prote
812	4	26.7	110	2	S60591	tg light chain var
813	4	26.7	110	2	S55879	ribosomal protein
814	4	26.7	110	2	S55877	ribosomal protein
815	4	26.7	110	2	JN0329	ethidium efflux pr
816	4	26.7	110	2	F90830	integral membrane
817	4	26.7	110	2	C85688	hypothetical prote
818	4	26.7	110	2	C71134	hypothetical prote
819	4	26.7	110	2	T47180	hypothetical prote
820	4	26.7	111	2	JC1110	ferredoxin [2Fe-2S
821	4	26.7	111	2	AC0352	ferredoxin family
822	4	26.7	111	2	A10833	ferredoxin [import
823	4	26.7	111	2	C85897	[2FE-2S] ferredoxi
824	4	26.7	111	2	G91052	[2FE-2S] ferredoxi
825	4	26.7	111	2	G72745	hypothetical prote
826	4	26.7	112	2	S34818	nifW protein - Rho
827	4	26.7	112	2	A84369	hypothetical prote
828	4	26.7	112	2	C98126	conserved hypotet
829	4	26.7	113	2	T44286	ferredoxin [2Fe-2S
830	4	26.7	113	2	P00275	polyprotein - hepa
831	4	26.7	113	2	P00276	polyprotein - hepa
832	4	26.7	113	2	F97624	csaa protein [impo

833	4	26.7	113	2	AG2847	secretion chaparon
834	4	26.7	113	2	S66709	probable membrane
835	4	26.7	113	2	AG0396	hypothetical prote
836	4	26.7	113	2	H72528	hypothetical prote
837	4	26.7	113	2	T15547	hypothetical prote
838	4	26.7	113	2	A10723	conserved hypothet
839	4	26.7	113	2	I40399	flagellar protein
840	4	26.7	113	2	S23653	sensorin A - Calif
841	4	26.7	113	2	T12933	hypothetical prote
842	4	26.7	113	2	H71182	hypothetical prote
843	4	26.7	114	1	R1D1S1	somatosatin-14 pr
844	4	26.7	114	1	GMPCP	cholecystokinin pr
845	4	26.7	114	1	R5HS2H	ribosomal protein
846	4	26.7	114	2	I50798	preprosomatostatin
847	4	26.7	114	2	F84266	50S ribosomal prot
848	4	26.7	114	2	A55010	neutrophil-activat
849	4	26.7	114	2	S09220	capsid protein - d
850	4	26.7	114	2	B75117	hypothetical prote
851	4	26.7	114	2	T17929	hypothetical prote
852	4	26.7	114	2	C75555	hypothetical prote
853	4	26.7	114	2	B98062	hypothetical prote
854	4	26.7	114	2	T50948	hypothetical prote
855	4	26.7	115	1	GMHCP	cholecystokinin pr
856	4	26.7	115	2	I51064	somatosatin II pr
857	4	26.7	115	2	A48318	cholecystokinin pr
858	4	26.7	115	2	B75567	hypothetical prote
859	4	26.7	115	2	E90443	hypothetical prote
860	4	26.7	115	2	T14461	non intermediate f
861	4	26.7	115	2	T50477	hypothetical prote
862	4	26.7	115	2	T33703	hypothetical prote
863	4	26.7	115	2	G72493	hypothetical prote
864	4	26.7	116	2	S78094	endocuticular prot
865	4	26.7	116	2	PC4092	alanine carrier pr
866	4	26.7	116	2	F81796	hypothetical inner
867	4	26.7	116	2	D95280	hypothetical prote
868	4	26.7	116	2	D95385	protein [imported
869	4	26.7	116	2	D72560	hypothetical prote
870	4	26.7	117	1	ERAD1	early E3 13K glyco
871	4	26.7	117	2	E83526	probable arsenate
872	4	26.7	117	2	C81247	arsenate reductase
873	4	26.7	117	2	S73025	hypothetical prote
874	4	26.7	117	2	C59092	hypothetical prote
875	4	26.7	117	2	F72536	hypothetical prote
876	4	26.7	117	2	T49497	hypothetical prote
877	4	26.7	117	2	B83187	hypothetical prote
878	4	26.7	117	2	B86862	hypothetical prote
879	4	26.7	117	2	C72582	hypothetical prote
880	4	26.7	118	1	PSSMK3	phospholipase A2 (
881	4	26.7	118	2	H34860	C-type natriuretic
882	4	26.7	118	2	B54119	hypothetical prote
883	4	26.7	118	2	A75108	hypothetical prote
884	4	26.7	118	2	C84859	hypothetical prote
885	4	26.7	118	2	B86593	hypothetical prote
886	4	26.7	118	2	G95856	hypothetical prote
887	4	26.7	118	2	D72609	hypothetical prote
888	4	26.7	118	2	H72644	probable arabinoga
889	4	26.7	119	2	AH1817	single-stranded DN
890	4	26.7	119	2	T35859	hypothetical prote
891	4	26.7	119	2	S74932	hypothetical prote
892	4	26.7	119	2	F49804	hypothetical prote
893	4	26.7	119	2	A72651	hypothetical prote
894	4	26.7	119	2	J01327	anther-specific pr
895	4	26.7	119	2	E64940	hypothetical prote
896	4	26.7	119	2	B90942	hypothetical prote
897	4	26.7	119	2	F85790	hypothetical prote
898	4	26.7	119	2	S14292	transcription acti
899	4	26.7	120	2	AC0714	conserved hypothet
900	4	26.7	120	2	T02586	hypothetical prote
901	4	26.7	120	2	F81882	hypothetical prote
902	4	26.7	120	2	G86402	hypothetical prote
903	4	26.7	121	1	C69190	conserved hypothet
904	4	26.7	121	2	AB2314	ferredoxin-thiore
905	4	26.7	121	2	JC1138	xylokinase (EC 2
906	4	26.7	121	2	G71566	probable l19 ribos
907	4	26.7	121	2	G72768	hypothetical prote
908	4	26.7	121	2	B70767	hypothetical prote
909	4	26.7	121	2	S58088	sex-determining pr
910	4	26.7	121	2	S43481	matutase homolog -
911	4	26.7	121	2	G83684	hexon-associated p
912	4	26.7	122	1	SXAD9T	aldehyde dehydroge
913	4	26.7	122	2	T14821	hypothetical prote
914	4	26.7	122	2	B72637	conserved hypothet
915	4	26.7	122	2	H72279	galanin precursor
916	4	26.7	123	1	RHHUN	galanin precursor
917	4	26.7	123	1	RHHOG	galanin precursor
918	4	26.7	123	1	RHPGN	galanin precursor
919	4	26.7	123	2	C61545	plasmin (EC 3.4.21
920	4	26.7	123	2	T43824	hypothetical prote
921	4	26.7	123	2	D84322	hypothetical prote
922	4	26.7	123	2	T20354	hypothetical prote
923	4	26.7	123	2	G75336	hypothetical prote
924	4	26.7	123	2	C71423	hypothetical prote
925	4	26.7	123	2	B87355	hypothetical prote
926	4	26.7	123	2	AH2525	hypothetical prote
927	4	26.7	123	2	C90432	conserved hypothet
928	4	26.7	123	2	D87238	hypothetical prote
929	4	26.7	123	2	G84592	hypothetical prote
930	4	26.7	123	2	G70969	hypothetical prote
931	4	26.7	123	2	C69282	conserved hypothet
932	4	26.7	123	2	T20279	hypothetical prote
933	4	26.7	124	1	RHRN	galanin precursor
934	4	26.7	124	2	JC5480	galanin precursor
935	4	26.7	124	2	A60218	galanin precursor
936	4	26.7	124	2	B83335	14k brain-specific
937	4	26.7	124	2	A70861	conserved hypothet
938	4	26.7	124	2	A72629	hypothetical prote
939	4	26.7	125	2	G47021	hypothetical prote
940	4	26.7	125	2	C32880	peptic enzyme secr
941	4	26.7	125	2	T36939	puls protein precu
942	4	26.7	125	2	T16848	hypothetical prote
943	4	26.7	125	2	F82974	hypothetical prote
944	4	26.7	125	2	B86051	hypothetical prote
945	4	26.7	125	2	C87560	hypothetical prote
946	4	26.7	125	2	E90350	hypothetical prote
947	4	26.7	125	2	E97414	hypothetical prote
948	4	26.7	125	2	G83555	hypothetical prote
949	4	26.7	126	1	AWHUC	natriuretic peptid
950	4	26.7	126	1	A36155	natriuretic peptid
951	4	26.7	126	1	H69178	conserved hypothet
952	4	26.7	126	2	S12988	brain natriuretic
953	4	26.7	126	2	A55688	natriuretic peptid
954	4	26.7	126	2	S29303	hypothetical prote
955	4	26.7	126	2	AF0356	holo-[acyl]-carrier
956	4	26.7	126	2	S65805	transcription acti
957	4	26.7	126	2	T03813	hypothetical prote
958	4	26.7	126	2	S64131	hypothetical prote
959	4	26.7	126	2	H72486	hypothetical prote
960	4	26.7	126	2	S74990	hypothetical prote
961	4	26.7	126	2	F89124	protein K07C11.10
962	4	26.7	126	2	AG2412	hypothetical prote
963	4	26.7	127	2	A70200	ribosome-binding f
964	4	26.7	127	2	D54759	terminal oxidase (
965	4	26.7	127	2	T11004	MGCL36.12 protein
966	4	26.7	127	2	D75595	hypothetical prote
967	4	26.7	127	2	F69070	conserved hypothet
968	4	26.7	127	2	S53573	conserved hypothet
969	4	26.7	127	2	G82084	probable membrane
970	4	26.7	128	2	A33083	hypothetical prote
971	4	26.7	128	2	S33632	pancreatic ribonuc
972	4	26.7	128	2	B75531	l1r1 protein - ric
973	4	26.7	128	2	S53584	hypothetical prote
974	4	26.7	128	2	AE2364	probable membrane
975	4	26.7	129	2	F81310	hypothetical prote
976	4	26.7	129	2	F61187	alpha-7A integrin
977	4	26.7	129	2	S69862	hypothetical prote
978	4	26.7	129	2	T36732	hypothetical prote

979 4 26.7 129 2 T36916 hypothetical prote  
980 4 26.7 129 2 T32313 hypothetical prote  
981 4 26.7 129 2 H86673 single-strand bind  
982 4 26.7 129 2 H72806 probable DNA prima  
983 4 26.7 129 2 A72606 hypothetical prote  
984 4 26.7 129 2 A97553 specialized acyl c  
985 4 26.7 130 2 S30534 ig heavy chain V r  
986 4 26.7 130 2 S31003 gene 58 protein -  
987 4 26.7 130 2 H83150 hypothetical prote  
988 4 26.7 130 2 T22920 hypothetical prote  
989 4 26.7 130 2 T51879 hypothetical prote  
990 4 26.7 130 2 E72697 hypothetical prote  
991 4 26.7 130 2 A97549 hypothetical prote  
992 4 26.7 130 2 AG2768 conserved hypotet  
993 4 26.7 131 2 E25733 T-cell receptor al  
994 4 26.7 131 2 B72557 probable ribosomal  
995 4 26.7 131 2 T11557 tat protein - simi  
996 4 26.7 131 2 S45059 AC1 protein (clone  
997 4 26.7 131 2 AC0747 conserved hypotet  
998 4 26.7 131 2 A69424 hypothetical prote  
999 4 26.7 131 2 E70920 probable moab prot  
1000 4 26.7 131 2 B75277 hypothetical prote

## ALIGNMENTS

## RESULT 1

S63683

C:Species: Homo sapiens (man)

C&gt;Date: 20-Jul-1996 #sequence\_revision 13-Mar-1997 #text\_change 29-Sep-1999

C:Accession: S63683

R:Sahara, N.; Yahagi, Y.; Takagi, H.; Kondo, T.; Okochi, M.; Usami, M.; Shirasawa, T.; M  
FEBS Lett. 381, 7-11, 1996

A&gt;Title: Identification and characterization of presenilin 1-467, I-463 and I-374.

A:Reference number: S63683; MUID:96193901; PMID:8641442

A:Accession: S63683

A&gt;Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-463 &lt;SAH&gt;

A:Cross-references: EMBL:040379; NID:q1244637; PIDN:AAB05894.1; PID:q1244638

C:Superfamily: Presenilin

Query Match 100.0%; Score 15; DB 2; Length 467;

Best Local Similarity 100.0%; Pred. No. 2.7e-09;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPRTSPESRAA 15

DB 342 SHLGPRTSPESRAA 356

## RESULT 2

S58396

C:Species: Homo sapiens (man)

C&gt;Date: 29-Jan-1998 #sequence\_revision 13-Feb-1998 #text\_change 29-Sep-1999

C:Accession: S58396; S71401; S71402

R:Sherrington, R.; Rogeev, E.I.; Liang, Y.; Rogeeva, E.A.; Levesque, G.; Ikeda, M.; Chl  
ero, I.; Pluessl, L.; Nee, L.; Chumakov, I.; Pollen, D.; Brookes, A.; Sansseau, P.; Polli  
E.; Rommens, J.M.; St George-Hyslop, P.H.

A&gt;Title: Cloning of a gene bearing missense mutations in early-onset familial Alzheimer

A:Reference number: 158095; MUID:95319502; PMID:7596406

A:Accession: S58396

A:Molecule type: mRNA

A:Residues: 1-467 &lt;SHS&gt;

A:Cross-references: EMBL:L42110; NID:9904118; PIDN:AAB64416.1; PID:9904119

R:Vidal, R.; Ghiso, J.; Wisniewski, T.; Frangione, B.

FEBS Lett. 393, 19-23, 1996

A&gt;Title: Alzheimer's presenilin 1 gene expression in platelets and megakaryocytes. Ident

A:Reference number: S71401; MUID:96397521; PMID:8804415  
A:Accession: S71401  
A>Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 24-32;254-256,290-292;316-317,376-379 <VW>  
A:Experimental source: Dam megakaryotic cell line (ATCC CRL-9792) and platelets  
C:Genetics:  
A:Gene: GDB:PSEN1; AD3; FAD; S182; PS1  
A:Cross-references: GDB:135682; OMIM:104311  
A:Map position: 14q24.3-14q24.3  
C:Superfamily: Presenilin  
C:Keywords: alternative splicing; Alzheimer's disease; glycoprotein; transmembrane pr  
F:82-100/Domain: transmembrane #status predicted <TM1>  
F:133-154/Domain: transmembrane #status predicted <TM2>  
F:164-185/Domain: transmembrane #status predicted <TM3>  
F:195-213/Domain: transmembrane #status predicted <TM4>  
F:221-238/Domain: transmembrane #status predicted <TM5>  
F:244-264/Domain: transmembrane #status predicted <TM6>  
F:281-301/Domain: transmembrane #status predicted <TM7>  
F:408-428/Domain: transmembrane #status predicted <TM8>  
F:433-453/Domain: transmembrane #status predicted <TM9>  
F:279,405/Binding site: carbohydrate (asn) (covalent) #status predicted

OY 1 SHLGPRTSPESRAA 15  
DB 346 SHLGPRTSPESRAA 360

## RESULT 3

I78388

C:Species: Mus musculus (house mouse)

C&gt;Date: 27-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 29-Sep-1999

C:Accession: I78388

R:Sherrington, R.; Rogeev, E.I.; Liang, Y.; Rogeeva, E.A.; Levesque, G.; Ikeda, M.; C  
ero, I.; Pluessl, L.; Nee, L.; Chumakov, I.; Pollen, D.; Brookes, A.; Sansseau, P.; Po  
E.; Rommens, J.M.; St George-Hyslop, P.H.

A&gt;Title: Cloning of a gene bearing missense mutations in early-onset familial Alzheim

A:Reference number: 158095; MUID:95319502; PMID:7596406

A:Accession: I78388

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-467 &lt;RES&gt;

A:Cross-references: GB:L42177; NID:9904129; PIDN:AAC42094.1; PID:9904130

C:Superfamily: Presenilin

Query Match 100.0%; Score 15; DB 2; Length 467;

Best Local Similarity 100.0%; Pred. No. 2.7e-09;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPRTSPESRAA 15

DB 346 SHLGPRTSPESRAA 360

## RESULT 4

H75253

C:Species: Deinococcus radiodurans (strain R1)

C&gt;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000

C:Accession: H75253

R:White, O.; Eissen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J  
M.; Shen, M.; Vamathevan, J.T.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.;  
Science 286, 1571-1577, 1999

A&gt;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: H75253

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-347 <MW>  
A:Cross-references: GB:AE002089; GB:AE000513; NID:g6460427; PIDN:AAF12144.1; PID:g646043  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR2600  
A:Map position: 1

Query Match 53.3%; Score 8; DB 2; Length 347;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSPESRA 14  
|||||||  
DB 274 RSPESRA 281

RESULT 5  
F87541  
hypothetical protein CC2359 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 17-May-2002  
C:Accession: F87541  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete genome sequence of Caulobacter crescentus.  
A:Reference number: AB7249; MUID:21173698; PMID:11259647  
A:Accession: F87541  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-327 <STO>  
A:Cross-references: GB:AE005673; NID:g13423888; PIDN:AAK24330.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC2359  
C:Superfamily: N4-(beta-N-acetylglucosaminyl)-L-asparaginase

Query Match 46.7%; Score 7; DB 2; Length 327;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRAA 15  
|||||||  
DB 139 TPESRAA 145

RESULT 6  
E95911  
probable transcription regulator, LacI family protein [imported] - Sinorhizobium meliloti

C:Species: Sinorhizobium meliloti  
C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C:Accession: E95911

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A:Title: The complete sequence of the 1.683-kb psymb megaplasmid from the N2-fixing endo  
A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: E95911  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-343 <KUR>

A:Cross-references: GB:AL591985; PIDN:CAC48957.1; PID:g15140442; GSPDB:GN00167

A:Experimental source: strain 1021, megaplasmid pSymb  
R:Gallbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Welis, D.H.; Wong, K.; Yeh, K.  
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation

C:Genetics:  
A:Gene: SMB20817  
A:Genome: plasmid

Query Match 46.7%; Score 7; DB 2; Length 343;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRAA 15  
|||||||  
DB 285 TPESRAA 291

RESULT 7

JC5081  
presenilin 1 protein isoform 463 - lesser mouse lemur

C:Species: Microcebus murinus (lesser mouse lemur)  
C:Date: 31-Jan-1997 #sequence\_revision 31-Jan-1997 #text\_change 13-Sep-1998  
C:Accession: JC5081  
R:Calenda, A.; Mestre-Frances, N.; Czech, C.; Pradier, L.; Petter, A.; Bons, N.; Bell  
Biochem. Biophys. Res. Commun. 228, 430-439, 1996  
A:Title: Molecular cloning, sequencing, and brain expression of the presenilin 1 gene  
A:Reference number: JC5080; MUID:97079199; PMID:8920931  
A:Contents: brain

A:Accession: JC5081  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-463 <CAL>

A:Cross-references: EMBL:271333  
C:Comment: This protein is an intermembrane protein with seven transmembrane domains.  
C:Genetics:  
A:Gene: psi

A:Map position: 14  
C:Superfamily: presenilin  
C:Keywords: transmembrane protein  
F:78-96/Domain: transmembrane #status predicted <TM1>  
F:129-150/Domain: transmembrane #status predicted <TM2>  
F:160-181/Domain: transmembrane #status predicted <TM3>  
F:191-209/Domain: transmembrane #status predicted <TM4>  
F:217-234/Domain: transmembrane #status predicted <TM5>  
F:240-257/Domain: transmembrane #status predicted <TM6>  
F:404-424/Domain: transmembrane #status predicted <TM7>

Query Match 46.7%; Score 7; DB 2; Length 463;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRST 9  
|||||||  
DB 344 LGPHRST 350

RESULT 8  
JC5080

presenilin 1 protein isoform 467 - lesser mouse lemur  
C:Species: Microcebus murinus (lesser mouse lemur)  
C:Date: 31-Jan-1997 #sequence\_revision 31-Jan-1997 #text\_change 20-Jun-2000  
C:Accession: JC5080

R:Calenda, A.; Mestre-Frances, N.; Czech, C.; Pradier, L.; Petter, A.; Bons, N.; Bell  
Biochem. Biophys. Res. Commun. 228, 430-439, 1996  
A:Title: Molecular cloning, sequencing, and brain expression of the presenilin 1 gene  
A:Reference number: JC5080; MUID:97079199; PMID:8920931

A:Accession: JC5080  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-467 <CAL>

A:Cross-references: EMBL:271333; NID:g1707591; PIDN:CAA5930.1; PID:g1707592  
A:Experimental source: brain

C:Comment: This protein is an intergal membrane protein with seven transmembrane dom  
C:Genetics:  
A:Gene: psi  
A:Map position: 14  
C:Superfamily: presenilin

C:Keywords: transmembrane protein  
F:92-100/Domain: transmembrane #status predicted <TM1>  
F:133-154/Domain: transmembrane #status predicted <TM2>  
F:164-185/Domain: transmembrane #status predicted <TM3>  
F:195-213/Domain: transmembrane #status predicted <TM4>  
F:221-238/Domain: transmembrane #status predicted <TM5>  
F:244-261/Domain: transmembrane #status predicted <TM6>  
F:408-428/Domain: transmembrane #status predicted <TM7>

Query Match 46.7%; Score 7; DB 2; Length 467;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRST 9  
Db 348 LGPHRST 354

## RESULT 9

S73457  
methionine-tRNA ligase (EC 6.1.1.10) mets - Mycoplasma pneumoniae (strain ATCC 29342)

N:Alternate names: hypothetical protein B01\_orf512; methionyl-tRNA synthetase mets

C:Species: Mycoplasma pneumoniae  
C:Accession: S73457  
C:Date: 26-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 03-Jun-2002

R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.

Nucleic Acids Res. 24, 4420-4449, 1996

A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae

A:Reference number: S73327; MUID:97105885; PMID:8948633

A:Accession: S73457  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA  
A:Residues: 1-512 <HIM>

A:Cross-references: EMBL:AE000015; GB:U00089; NID:g1673779; PIDN:AAB95779.1; PID:g1673779

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996

C:Genetics:  
A:Gene: mets  
A:Genetic code: SGC3

C:Superfamily: methionine-tRNA synthetase; ligase; protein biosynthesis

C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 46.7%; Score 7; DB 2; Length 512;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRAA 15  
Db 180 TPESRAA 186

## RESULT 10

F45714  
probable regulatory protein p13 II, alternative splice form - human T-cell lymphotropic

C:Species: human T-cell lymphotropic virus type 1, HTLV-1  
C:Date: 22-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 23-May-1997

C:Accession: F45714  
R:Gessain, A.; Boeri, E.; Yanagihara, R.; Gallo, R.C.; Franchini, G.

J. Virol. 67, 1015-1023, 1993

A:Title: Complete nucleotide sequence of a highly divergent human T-cell leukemia (lymph)

from other geographical regions.

A:Reference number: A45714; MUID:93124536; PMID:8419636

A:Contents: HTLV-IMEL5  
A:Accession: F45714

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid  
A:Residues: 1-87 <GES>  
A:Note: sequence extracted from NCBI backbone (NCBIP:122473)

Query Match 40.0%; Score 6; DB 2; Length 87;  
Best Local Similarity 100.0%; Pred. No. 7;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPHR 7  
Db 58 HLGPHR 63

## RESULT 11

D61547  
hypothetical protein II (px region) - human T-cell lymphotropic virus type 1 (isolate

C:Species: human T-cell lymphotropic virus type 1, HTLV-1  
C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 08-Oct-1999

C:Accession: D61547  
R:Tsujimoto, A.; Teruuchi, T.; Imamura, J.; Shimotohno, K.; Miyoshi, I.; Miya, M.

Mol. Biol. Med. 5, 29-42, 1988

A:Title: Nucleotide sequence analysis of a provirus derived from HTLV-1-associated my

A:Reference number: A61547; MUID:88232270; PMID:2897612

A:Accession: D61547  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-87 <TSU>

A:Cross-references: GB:M37301; NID:g541634; PIDN:AAA45391.1; PID:g541638

Query Match 40.0%; Score 6; DB 2; Length 87;  
Best Local Similarity 100.0%; Pred. No. 7;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPHR 7  
Db 58 HLGPHR 63

## RESULT 12

D95289  
hypothetical protein Sma0412 [imported] - Sinorhizobium meliloti (strain 1021) megapl

C:Species: Sinorhizobium meliloti  
C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001

C:Accession: D95289  
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B

.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meli

A:Reference number: A95262; MUID:21396509; PMID:11481432

A:Accession: D95289  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-138 <KUR>

A:Cross-references: GB:AE006469; PIDN:AAK64878.1; PID:g14523295; GSPDB:GN00165

A:Experimental source: strain 1021, megaplasmid PSymA

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, A.P.; Ampe, F.; Barloy-Hubl

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation  
C:Genetics:  
A:Gene: Sma0412  
A:Genome: plasmid

Query Match 40.0%; Score 6; DB 2; Length 138;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8  
Db 109 LGPHRS 114

## RESULT 13

T26722  
hypothetical protein Y39A1A.18 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans



C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C/Accession: T26722  
 R:Wall, M.  
 submitted to the EMBL Data Library, September 1998  
 A:Reference number: Z20257  
 A:Accession: T26722  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-139 <Wt>  
 A:Cross-references: EMBL:AL031633; PIDN:CAA21011.1; GSPDB:GN00021; CESP:Y39A1A.18  
 A:Experimental source: clone Y39A1A  
 C:Genetics:  
 A:Gene: CESP:Y39A1A.18  
 A:Map position: 3  
 A:Introns: 37/3; 79/3

Query Match  
 Best Local Similarity 40.0%; Score 6; DB 2; Length 139;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14  
 111111  
 DB 19 TPESRA 24

RESULT 14  
 E75257  
 molybdenum cofactor biosynthesis protein C - Deinococcus radiodurans (strain R1)  
 C:Species: Deinococcus radiodurans  
 C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
 C/Accession: E75257  
 R:White, O.; Eisen, J.A.; Weidberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567266  
 A:Accession: E75257  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-168 <Wt>  
 A:Cross-references: GB:AE002086; GB:AE000513; NID:g6460395; PIDN:AAF12111.1; PID:g646039  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR2571  
 A:Map position: 1  
 C:Superfamily: molybdenum cofactor biosynthesis protein C

Query Match  
 Best Local Similarity 40.0%; Score 6; DB 2; Length 168;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15  
 111111  
 DB 44 PESRAA 49

RESULT 15  
 E45714  
 probable regulatory function protein p30 II, alternative splice form - human T-cell lymph  
 C:Species: human T-cell lymphotropic virus type 1, HTLV-1  
 C>Date: 22-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 23-May-1997  
 C/Accession: E45714  
 R:Gessain, A.; Boert, E.; Yanagihara, R.; Gallo, R.C.; Franchini, G.  
 J. Virol. 67, 1015-1023, 1993  
 A:Title: Complete nucleotide sequence of a highly divergent human T-cell leukemia (lymph  
 from other geographical regions.  
 A:Reference number: A45714; MUID:93124536; PMID:8419636  
 A:Contents: HTLV-IMEL5  
 A:Accession: E45714  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 1-240 <GES>

A>Note: sequence extracted from NCBI backbone (NCBI:P.122472)

Query Match  
 Best Local Similarity 40.0%; Score 6; DB 2; Length 240;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPHR 7  
 111111  
 DB 211 HLGPHR 216

RESULT 16  
 C46181  
 pX-tax-orf II (alternatively spliced) - human T-cell lymphotropic virus type 1  
 C:Species: human T-cell lymphotropic virus type 1, HTLV-1  
 C>Date: 22-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 17-Nov-2000  
 C/Accession: C46181; D46181  
 R:Koralnik, I.J.; Gessain, A.; Klotman, M.E.; Lo Monaco, A.; Berneman, Z.N.; Franchin  
 Proc. Natl. Acad. Sci. U.S.A. 89, 8813-8817, 1992  
 A:Title: Protein isoforms encoded by the pX region of human T-cell leukemia/lymphotro  
 A:Reference number: A46181; MUID:92409607; PMID:1328697  
 A:Accession: C46181  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 1-241 <KOR>  
 A>Note: sequence extracted from NCBI backbone (NCBI:P.114307)  
 A:Accession: D46181  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 155-241 <KO2>  
 A>Note: sequence extracted from NCBI backbone (NCBI:P.114308)

Query Match  
 Best Local Similarity 40.0%; Score 6; DB 2; Length 241;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPHR 7  
 111111  
 DB 212 HLGPHR 217

RESULT 17  
 D96016  
 probable exodeoxyribonuclease III (EC 3.1.11.2) [imported] - Sinorhizobium meliloti  
 C:Species: Sinorhizobium meliloti  
 C>Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 14-Sep-2001  
 C/Accession: D96016  
 R:Finan, T.M.; Weidner, S.; Wong, K.; Bhramare, J.; Chalm, P.; Vorholter, F.J.; Hei  
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
 A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing  
 A:Reference number: A95842; MUID:21396508; PMID:11481431  
 A:Accession: D96016  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-257 <KUR>  
 A:Cross-references: GB:AL591985; PIDN:CAC49796.1; PID:g15141283; GSPDB:GN00167  
 A:Experimental source: strain 1021, megaplasmid pSymB  
 R:Gallbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hub  
 Pella, D.; Chalm, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.  
 Science 293, 668-672, 2001  
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau  
 hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Welis, D.H.; Wong, K.; Yeh,  
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
 A:Reference number: A96039; MUID:21368234; PMID:11474104  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: xtnA4; SMD20689  
 A:Accession: SMD20689  
 A:Status: preliminary  
 C:Superfamily: exodeoxyribonuclease III  
 C:Keywords: hydrolase

Query Match  
 40.0%; Score 6; DB 2; Length 257;

Best Local Similarity 100.0%; Pred. No. 19;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRAA 15  
|||||  
Db 171 PESRAA 176

RESULT 18  
F87498

exodeoxyribonuclease III [imported] - *Caulobacter crescentus*

C:Species: *Caulobacter crescentus*  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001

A:Accession: F87498  
R:NIEMAN, W.C.; FELDBLYUM, T.V.; PAULSEN, I.T.; NELSON, K.E.; EISEN, J.; HELDELBERG, J.  
B.; LAUB, M.T.; DEBOY, R.T.; DODSON, R.J.; DURKIN, A.S.; GWINN, M.L.; HAFL, D.H.; KOLON  
N, J.; EMOJAEVA, M.; WHITE, O.; SELZBERG, S.L.; SHAPIRO, L.; VENTER, J.C.; FRASER, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete genome sequence of *Caulobacter crescentus*.  
A:Reference number: A87249; MUID:21173698; PMID:1259647

A:Accession: F87498  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-262 <SNQ>

A:Cross-references: GB:AE005673; NID:q13423482; PIDN:AAK23986.1; GSPDB:GN00148

A:Gene: CC2011

C:Superfamily: exodeoxyribonuclease III

Query Match 40.0%; Score 6; DB 2; Length 262;  
Best Local Similarity 100.0%; Pred. No. 19;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRAA 15  
|||||  
Db 176 PESRAA 181

RESULT 19  
T37031

hypothetical protein SCJ12.12c - *Streptomyces coelicolor*

C:Species: *Streptomyces coelicolor*  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T37031  
R:MURPHY, L.; HARRIS, D.; THOMSON, N.R.; PARKHILL, J.; BARRELL, B.G.; RAJANDREAM, M.A.  
submitted to the EMBL Data Library, August 1999

A:Reference number: Z21619

A:Accession: T37031

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA  
A:Residues: 1-301 <MUR>

A:Cross-references: EMBL:AL109989; PIDN:CA53424.1; GSPDB:GN00070; SCODEB:SCJ12.12c

C:Genetics:  
A:Gene: SCODEB:SCJ12.12c

Query Match 40.0%; Score 6; DB 2; Length 301;  
Best Local Similarity 100.0%; Pred. No. 21;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRAA 15  
|||||  
Db 13 PESRAA 18

RESULT 20  
A82860

DNA replication and repair RecF protein XF0003 [imported] - *Xylella fastidiosa* (strain 9

C:Species: *Xylella fastidiosa*

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 17-Nov-2000

C:Accession: A82860  
R:anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: A82860  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-364 <SIM>  
A:Cross-references: GB:AE003855; GB:AE003849; NID:g9104760; PIDN:AAE82816.1; GSPDB:GN  
A:Experimental source: strain 9a5c  
R:SIMPSON, A.J.G.; REINACH, F.C.; ARRUDA, P.; ABREU, F.A.; ACENCIO, M.; ALVARENGA, R.  
BRIONES, M.R.S.; BUENO, M.R.P.; CAMARGO, A.A.; CAMARGO, L.E.A.; CARRARO, D.M.; CARRER  
AS-NEITO, E.; DOCEMA, C.; EL-DORRY, H.; FACINCANI, A.P.; FERREIRA, A.J.S.  
submitted to GenBank, June 2000

A:Authors: FERREIRA, V.C.A.; FERRO, J.A.; FRAGA, J.S.; FRANCA, S.C.; FRANCO, M.C.; FR  
J.D.; JUNQUEIRA, M.L.; KEMPER, E.L.; KITAJIMA, J.P.; KRIEGER, J.E.; KURAMAE, E.E.; LA  
CHADO, M.A.; MADEIRA, A.M.B.N.; MADEIRA, H.M.F.; MARINO, C.L.; MARGUES, M.V.; MARTINS  
A:Authors: MARTINS, E.M.F.; MATSUKUMA, A.Y.; MENCK, C.F.M.; MIRACCA, E.C.; MIYAKI, C.  
F.G.; NUNES, L.R.; OLIVEIRA, M.A.; DE OLIVEIRA, M.C.; DE OLIVEIRA, R.C.; PALMERI,  
RODRIGUES, V.; ROSA, A.J. de M.; de ROSA JR., V.E.; de SA, R.G.; SANTILLI, R.V.; SAVA  
A:Authors: da SILVA, A.C.R.; da SILVA, F.R.; da SILVA, A.M.; SILVA JR., W.A.; da SILV  
M.; TSUNAKO, M.H.; VALLADA, H.; VAN SLUYS, M.A.; VERJOVSKI-ALMEIDA, S.; VETTORE, A.L.  
A:Reference number: A59328

A:Contents: annotation

C:Genetics:  
A:Gene: XF0003

C:Superfamily: recF protein

Query Match 40.0%; Score 6; DB 2; Length 364;  
Best Local Similarity 100.0%; Pred. No. 25;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHRS 8  
|||||  
Db 254 LGPHRS 259

RESULT 21  
S40693

opsin rh3 - fruit fly (*Drosophila pseudoobscura*)

C:Species: *Drosophila pseudoobscura*  
C:Date: 19-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 13-Aug-1999

C:Accession: S40693; S65548; S24607  
R:CARULLI, J.P.; HARTEL, D.L.  
Genetics 132, 193-204, 1992

A:Title: Variable rates of evolution among *Drosophila* opsin genes.

A:Reference number: S40691; MUID:93012921; PMID:1398053

A:Accession: S40693

A:Molecule type: DNA  
A:Residues: 1-382 <CAR>

A:Cross-references: EMBL:X65879

A:Note: the authors translated the codon GCA for residue 25 as Gly

R:CARULLI, J.P.  
submitted to the EMBL Data Library, November 1992

A:Reference number: S65546

A:Accession: S65548

A:Molecule type: DNA  
A:Residues: 1-123, 'L', 125-241, 'Q', 243-254, 'R', 256-382 <CAW>

A:Cross-references: EMBL:X65879; NID:g9081; PIDN:CAA46710.1; PID:g9082

C:Genetics:  
A:Gene: FlyBase:Dpse/Rh3

A:Cross-references: FlyBase:FBgn0012709

C:Superfamily: vertebrate rhodopsin

Query Match 40.0%; Score 6; DB 2; Length 382;  
Best Local Similarity 100.0%; Pred. No. 26;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRAA 15  
|||||  
Db 362 PESRAA 367

RESULT 22

F69834  
multidrug-efflux transporter homolog yhjO - *Bacillus subtilis*  
C:Species: *Bacillus subtilis*  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
C:Accession: F69834  
R:Kunst, F.; Ogawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berlet  
C.; Bron, S.; Brulliet, S.; Busch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fuma, S.; Galizzi, A.; Gallier  
lech, J.; Harwood, C.R.; Henaute, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinis,  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
Y, M.; Ogawa, K.; Ogawara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A:Authors: Schleich, S.; Schroeter, R.; Scottone, F.; Sekiguchi, J.; Sekowska, A.; Serot  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpiltra, P.; Tognoni, A.; Tosato, V.; Uchlyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: F69834  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Residues: 1-401 <KUN>  
A:Molecule type: DNA  
A:Cross-references: GB:299109; GB:AL009126; NID:g2633260; PIDN:CAM12898.1; PID:g2633394  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yhjO  
C:Superfamily: tetracycline resistance protein

Query Match      40.0%    Score 6;    DB 2;    Length 401;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches    6;    Conservative    0;    Mismatches    0;    Indels    0;    Gaps    0;

OY    8 STPESR 13  
      |||||  
Db    131 STPESR 136

RESULT 23  
T34535  
hypothetical protein DKFZp434H0717.1 - human (fragment)  
C:Species: *Homo sapiens (man)*  
C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T34535  
R:Bioecker, H.; Boecker, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, October 1999  
A:Reference number: 221539  
A:Accession: T34535  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-407 <BLO>  
A:Cross-references: EMBL:AL122102  
A:Experimental source: adult testis; clone DKFZp434H0717  
C:Genetics:  
A:Note: DKFZp434H0717.1

Query Match      40.0%    Score 6;    DB 2;    Length 407;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches    6;    Conservative    0;    Mismatches    0;    Indels    0;    Gaps    0;

OY    7 RSTPES 12  
      |||||  
Db    23 RSTPES 28

RESULT 24  
D87297  
conserved hypothetical protein CC0389 [Imported] - *Caulobacter crescentus*  
C:Species: *Caulobacter crescentus*  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: D87297

R:Nieman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,  
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: D87297  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-408 <STO>  
A:Cross-references: GB:AE005673; NID:913421548; PIDN:AAK22376.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC0389

Query Match      40.0%    Score 6;    DB 2;    Length 408;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches    6;    Conservative    0;    Mismatches    0;    Indels    0;    Gaps    0;

OY    10 PESRAA 15  
      |||||  
Db    117 PESRAA 122

RESULT 25  
C95248  
histidyl-tRNA synthetase [Imported] - *Streptococcus pneumoniae* (strain TIGR4)  
C:Species: *Streptococcus pneumoniae*  
C>Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 24-Aug-2001  
C:Accession: C95248  
R:Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H  
nson, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzappl  
nson, T.; Hickay, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris  
A:Title: Complete Genome Sequence of a virulent isolate of *Streptococcus pneumoniae*.  
A:Reference number: A95000; MUID:21357209; PMID:11463916  
A:Accession: C95248  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-429 <KUN>  
A:Cross-references: GB:AE005672; PIDN:AAK76180.1; PID:914973634; GSPDB:GN00164; TIGR:  
A:Experimental source: strain TIGR4  
C:Genetics:  
A:Gene: SP2121  
C:Superfamily: histidine-tRNA ligase; histidine-tRNA ligase homology

Query Match      40.0%    Score 6;    DB 2;    Length 429;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches    6;    Conservative    0;    Mismatches    0;    Indels    0;    Gaps    0;

OY    10 PESRAA 15  
      |||||  
Db    171 PESRAA 176

RESULT 26  
H98112  
histidine-tRNA ligase (EC 6.1.1.21) [Imported] - *Streptococcus pneumoniae* (strain R6)  
C:Species: *Streptococcus pneumoniae*  
C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 03-Jun-2002  
C:Accession: H98112  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAlhnen, S.  
y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.  
A:Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: H98112  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-429 <KUN>  
A:Cross-references: GB:AE007317; PIDN:AAU00733.1; PID:915459628; GSPDB:GN00174

C:Genetics:  
A:Gene: hist  
C:Superfamily: histidine-trNA ligase; histidine-trNA ligase homology  
C:Keywords: ligase

Query Match 40.0%; Score 6; DB 2; Length 429;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 10 PESRAA 15  
|||||  
DB 171 PESRAA 176

RESULT 27  
JC5390  
Presentin-alpha - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C>Date: 04-Jun-1997 #sequence\_revision 18-Jul-1997 #text\_change 20-Jun-2000  
C:Accession: JC5390  
R:Tsujimura, A.; Yasojima, K.; Hashimoto-Gotoh, T.  
Biochem. Biophys. Res. Commun. 231, 392-396, 1997  
A:Title: Cloning of Xenopus presentin-alpha and -beta cDNAs and their differential exp  
A:Reference number: JC5390; MUID:9723465; PMID:9070286  
A:Accession: JC5390  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-433 <TSU>  
A:Cross-references: DDBJ:DB4427; NID:g1944353; PIDN:BA19570.1; PID:g1944354  
A:Experimental source: brain  
C:Comment: This protein plays a role in negative regulation of apoptotic cascades during  
C:Superfamily: presentin  
F:48-66/Domain: transmembrane #status predicted <TM1>  
F:99-119/Domain: transmembrane #status predicted <TM2>  
F:130-149/Domain: transmembrane #status predicted <TM3>  
F:161-178/Domain: transmembrane #status predicted <TM4>  
F:187-203/Domain: transmembrane #status predicted <TM5>  
F:210-227/Domain: transmembrane #status predicted <TM6>  
F:314-394/Domain: transmembrane #status predicted <TM7>

Query Match 40.0%; Score 6; DB 2; Length 433;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 8 STPESR 13  
|||||  
DB 319 STPESR 324

RESULT 28  
S31927  
drop9 protein - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 10-Sep-1997  
C:Accession: S31927  
R:Kirsch, K.  
submitted to the EMBL Data Library, February 1993  
A:Reference number: S31927  
A:Accession: S31927  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-475 <KIR>  
A:Cross-references: EMBL:X70991; NID:g38459; PID:g38460

Query Match 40.0%; Score 6; DB 2; Length 475;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 RSTPES 12  
|||||  
DB 139 RSTPES 144

RESULT 29  
T03510  
hypothetical protein - Rhodobacter capsulatus  
C:Species: Rhodobacter capsulatus  
C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 08-Oct-1999  
C:Accession: T03510  
R:Vicek, C.; Paces, V.; Maltsev, N.; Paces, J.; Haselkorn, R.; Fonstein, M.  
Proc. Natl. Acad. Sci. U.S.A. 94, 9384-9388, 1997  
A:Title: Sequence of a 189-kb segment of the chromosome of Rhodobacter capsulatus S91  
A:Reference number: Z14955; MUID:97404404; PMID:9256491  
A:Accession: T03510  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-617 <VLC>  
A:Cross-references: EMBL:AF010496; NID:g3128256; PIDN:AC16163.1; PID:g3128311  
C:Genetics:  
A:Map position: 1

Query Match 40.0%; Score 6; DB 2; Length 617;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 10 PESRAA 15  
|||||  
DB 470 PESRAA 475

RESULT 30  
AE1895  
hypothetical protein all0711 [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp.  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
C:Accession: AE1895  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Itigun  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata  
DNA Res. 8, 205-213, 2001  
A:Title: Complete genomic sequence of the filamentous Nitrogen-fixing Cyanobacterium  
A:Reference number: AB1895; MUID:21595285; PMID:11759840  
A:Accession: AE1895  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-640 <RUR>  
A:Cross-references: GB:BA000019; PIDN:BA072668.1; PID:g17130056; GSPDB:GN00179  
C:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all0711

Query Match 40.0%; Score 6; DB 2; Length 640;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 10 PESRAA 15  
|||||  
DB 142 PESRAA 147

RESULT 31  
I64207  
preprotein translocase secA - Mycoplasma genitalium  
C:Species: Mycoplasma genitalium  
C>Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 17-Nov-2000  
C:Accession: I64207  
R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.  
M.; Fuhmann, J.; Nguyen, D.; Uitterlbeck, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick,  
C.A.; Venter, J.C.  
Science 270, 397-403, 1995  
A:Title: The minimal gene complement of Mycoplasma genitalium.  
A:Reference number: A64200; MUID:96026346; PMID:7569993  
A:Accession: I64207  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-806 <TIGR>

A:Cross-references: GB:039687; GB:L43967; NID:g3844663; PIDN:AAC71290.1; PID:g1045748; I  
A:Experimental source: strain G-37  
C:Comment: The "nucleotide-binding motif B" and "DEAD motif" features as annotated are c  
is adjacent to the identified motif and a third conserved motif is approximately 120-140  
C:Genetics:  
A:Genetic code: SGC3  
C:Superfamily: preprotein translocase secA  
C:Keywords: ATP; membrane-associated complex; P-loop; protein transport  
F:102-109/Region: nucleotide-binding motif A (P-loop) #status atypical  
F:205-210/Region: nucleotide-binding motif B  
F:209-212/Region: DEAD motif

Query Match 40.0%; Score 6; DB 2; Length 806;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 TPESRA 15  
|||||  
Db 255 TPESRA 260

## RESULT 32

RRVGT  
RNA-directed RNA polymerase (EC 2.7.7.48) - tomato bushy stunt virus (strain cherry)  
N:Alternate names: RNA nucleotidyltransferase (RNA-directed); RNA replicase  
N:Contains: 33k protein  
C:Species: tomato bushy stunt virus, TBSV  
C:Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 25-Oct-1996  
R:Accession: A5315  
R:Heurstein, P.O.; Knorr, D.A.; Hillman, B.I.; Morris, T.J.  
Virology 177, 141-151, 1990  
A:Title: The complete genome structure and synthesis of infectious RNA from clones of to  
A:Reference number: A5315; MUID:90281577; PMID:2353450  
A:Accession: A5315  
A:Molecule type: genomic RNA  
A:Residues: 1-817 <HEA>  
A:Cross-references: EMBL:M31019  
A:Note: readthrough of the terminator UAG occurs between codons AAA for 296-Lys and GGA  
C:Superfamily: carnation mottle virus RNA-directed RNA polymerase; barley yellow dwarf v  
C:Keywords: nucleotidyltransferase; RNA biosynthesis; RNA replication  
F:1-296/Product: 33k protein #status predicted <P33>  
F:527-694/Domain: barley yellow dwarf virus RNA-directed RNA polymerase homology <BYD>

Query Match 40.0%; Score 6; DB 1; Length 817;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14  
|||||  
Db 766 TPESRA 771

## RESULT 33

RRVGT  
RNA-directed RNA polymerase (EC 2.7.7.48) - Cymbidium ringspot virus  
N:Alternate names: RNA nucleotidyltransferase (RNA-directed); RNA replicase  
N:Contains: 33k protein  
C:Species: Cymbidium ringspot virus  
C:Date: 31-Mar-1990 #sequence\_revision 01-Dec-2000 #text\_change 01-Dec-2000  
R:Accession: S05456; J50268  
R:Griceco, F.; Burgan, J.; Russo, M.  
Nucleic Acids Res. 17, 6383, 1989  
A:Title: The nucleotide sequence of Cymbidium ringspot virus RNA.  
A:Reference number: J50268; MUID:8936663; PMID:2771646  
A:Accession: S05456  
A:Status: preliminary; translation not shown  
A:Molecule type: genomic RNA  
A:Residues: 1-818 <GRI>  
A:Cross-references: EMBL:X15511; NID:g59020; PIDN:CAB38439.1; PID:g4469160  
A:Accession: J50268  
A:Molecule type: genomic RNA  
A:Residues: 1-296, 298-483, 1, 485-818 <GR2>  
A:Cross-references: GB:X15511

A:Note: readthrough of the terminator TAG occurs between codons AAA for 296-Lys and G  
C:Superfamily: carnation mottle virus RNA-directed RNA polymerase; barley yellow dwar  
C:Keywords: nucleotidyltransferase; RNA biosynthesis; RNA replication  
F:1-296/Product: 33k protein #status predicted <TTP>  
F:528-695/Domain: barley yellow dwarf virus RNA-directed RNA polymerase homology <BYD

Query Match 40.0%; Score 6; DB 1; Length 818;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14  
|||||  
Db 767 TPESRA 772

## RESULT 34

S52717  
RNA-directed RNA polymerase (EC 2.7.7.48) - carnation Italian ringspot virus  
N:Alternate names: RNA replicase  
C:Species: carnation Italian ringspot virus  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 18-Jun-1999  
R:Accession: S52717  
R:Rubino, L.; Burgan, J.; Russo, M.  
submitted to the EMBL Data Library, March 1995  
A:Description: Molecular cloning and complete nucleotide sequence of carnation Italia  
A:Reference number: S52717  
A:Accession: S52717  
A:Molecule type: genomic RNA  
A:Residues: 1-851 <RUB>  
A:Cross-references: EMBL:X85215; NID:g755716; PIDN:CAAS9478.1; PID:g755718  
C:Superfamily: carnation mottle virus RNA-directed RNA polymerase; barley yellow dwar  
C:Keywords: nucleotidyltransferase; RNA biosynthesis; RNA replication  
F:561-728/Domain: barley yellow dwarf virus RNA-directed RNA polymerase homology <BYD

Query Match 40.0%; Score 6; DB 2; Length 851;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14  
|||||  
Db 800 TPESRA 805

## RESULT 35

VGBRB  
glycoprotein B precursor - Marek's disease virus (strain RB1B)  
C:Species: Marek's disease virus  
C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 16-Jun-2000  
R:Accession: A32402; B32402  
R:Ross, L.J.N.; Sanderson, M.; Scott, S.D.; Blms, M.M.; Doel, T.; Milne, B.  
J. Gen. Virol. 70, 1789-1804, 1989  
A:Title: Nucleotide sequence and characterization of the Marek's disease virus homolo  
A:Reference number: A32402; MUID:89293086; PMID:2544666  
A:Accession: A32402  
A:Molecule type: DNA  
A:Residues: 1-865 <ROS>  
A:Cross-references: GB:D13713; EMBL:D00506; NID:g221836; PIDN:BAA02866.1; PID:g221837  
A:Accession: B32402  
A:Molecule type: protein  
A:Residues: 250-271; 304-330 <ROS2>  
C:Superfamily: herpesvirus glycoprotein B  
C:Keywords: glycoprotein; transmembrane protein  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-865/Product: glycoprotein B #status predicted <GPB>  
F:709-728/Domain: transmembrane #status predicted <TM1>  
F:732-752/Domain: transmembrane #status predicted <TM2>  
F:27,184,332,364,406,425,631/Binding site: carbohydrate (Asn) (covalent) #status pred

Query Match 40.0%; Score 6; DB 1; Length 865;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14

Db 174 TPESRA 179

## RESULT 36

probable major core protein 113L - Molluscum contagiosum virus 1  
N:Alternate names: MC113L  
C:Species: Molluscum contagiosum virus 1  
C>Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 21-Jul-2000  
C:Accession: T30715  
R:Senkevitch, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Dairi, G.; Moss, B.  
S:Science 273, 813-816, 1996  
A:Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re  
A:Reference number: Z20876; MUID:96325459; PMID:8670425  
A:Accession: T30715  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-889 <SEN>  
A:Cross-references: EMBL:U60315; NID:q1491943; PID:AMC5241.1; PID:q1492056  
C:Genetics:  
A:Note: MC113L  
C:Superfamily: vaccinia virus major core protein p4a

Query Match 40.0%; Score 6; DB 2; Length 889;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14

Db 211 TPESRA 216

## RESULT 37

B88794  
protein K07F5.12a [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C:Accession: B88794  
R:anonymous, The C. elegans Sequencing Consortium.  
S:Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C\_ele  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: B88794  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-981 <STO>  
A:Cross-references: GB:chr\_IV; PIDN:CAA94287.1; PID:g4008375; GSPDB:GN00022; CESP:K07F5.  
C:Genetics:  
A:Gene: K07F5.12a  
A:Map position: 4

Query Match 40.0%; Score 6; DB 2; Length 981;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 PHSRP 10

Db 533 PHSRP 538

## RESULT 38

T30815  
platelet-derived growth factor receptor beta - Japanese pufferfish  
C:Species: Fugu rubripes (Japanese pufferfish)  
C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Jun-2000  
C:Accession: T30815  
R:How, G.F.; Venkatesh, B.; Brenner, S.  
S:Genome Res. 6, 1185-1191, 1996  
A:Title: Conserved linkage between the pufferfish (Fugu rubripes) and human genes for pla  
A:Reference number: Z20882; MUID:97129405; PMID:8973913

A:Accession: T30815  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1048 <HOW>  
A:Cross-references: EMBL:U63926; NID:q1752706; PID:q1752707; PIDN:AMC60062.1  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo

Query Match 40.0%; Score 6; DB 2; Length 1048;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 6

Db 652 SHLGP 657

## RESULT 39

PFMSRB  
platelet-derived growth factor receptor beta precursor - mouse  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)  
C:Species: Mus musculus (house mouse)  
C>Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 10-Sep-1999  
C:Accession: A25742  
R:Yarden, Y.; Escobedo, J.A.; Kuang, W.J.; Yang-Feng, T.L.; Daniel, T.O.; Tremble, P.  
S:Nature 323, 226-232, 1986  
A:Title: Structure of the receptor for platelet-derived growth factor helps define a  
A:Reference number: A25742; MUID:87014762; PMID:3020426  
A:Accession: A25742

A:Molecule type: mRNA  
A:Residues: 1-1098 <YAR>  
A:Cross-references: EMBL:X04367; NID:953618; PIDN:CAA27882.1; PID:953619  
A:Note: part of this sequence, including the amino end of the mature protein, was con  
C:Comment: The extracellular domain is predicted to include five immunoglobulin-like  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo  
C:Keywords: ATP; autophosphorylation; glycoprotein; growth factor receptor; heterodim  
F:1-31/Domain: signal sequence #status predicted <SIG>  
F:32-530/Product: platelet-derived growth factor receptor beta #status predicted <MA  
F:32-530/Domain: extracellular #status predicted <EXT>  
F:46-101/Domain: immunoglobulin homology <IMM1>  
F:141-191/Domain: immunoglobulin homology <IMM2>  
F:227-282/Domain: immunoglobulin homology <IMM3>  
F:428-509/Domain: immunoglobulin homology <IMM4>  
F:531-554/Domain: transmembrane #status predicted <TMN>  
F:555-1098/Domain: intracellular #status predicted <INT>  
F:557-964/Domain: protein kinase homology <KIN>  
F:605-613/Region: protein kinase ATP-binding motif  
F:44,88,102,214,291,306,353,370,444,467,478/Binding site: carbohydrate (Asn) (covalen  
F:53-99,148-189,234-290,435-507/Disulfide bonds: #status predicted  
F:633/Active site: Lys #status predicted  
F:856/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi

Query Match 40.0%; Score 6; DB 1; Length 1098;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 6

Db 655 SHLGP 660

## RESULT 40

PFHUGB  
platelet-derived growth factor receptor beta precursor - human  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)  
C:Species: Homo sapiens (hmn)  
C>Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 11-Jun-1999  
C:Accession: A28206; A31195; A38268; A31925; B31925  
R:Gronwald, R.G.K.; Grant, F.J.; Haldean, B.A.; Hart, C.E.; O'Hara, P.J.; Hagen, F.S  
S:Proc. Natl. Acad. Sci. U.S.A. 85, 3435-3439, 1988  
A:Title: Cloning and expression of a cDNA coding for the human platelet-derived growt  
A:Reference number: A28206; MUID:88217915; PMID:2835772  
A:Accession: A28206

A:Molecule type: mRNA  
A:Residues: 1-1106 <GRO>  
A:Cross-References: GB:J03278; NID:9189731; PIDN:AA60049.1; PID:9189732  
R:Classon-Welsh, L.; Eriksson, A.; Møren, A.; Severinsson, L.; Ek, B.; Oestman, A.; Bet  
Mol. Cell. Biol. 8, 3476-3486, 1988  
A:Title: cDNA cloning and expression of a human platelet-derived growth factor (PDGF) r  
A:Reference number: A31195; MUID:89096941; PMID:2850496  
A:Accession: A31195  
A:Molecule type: mRNA  
A:Residues: 1-240, 'D', 242-1106 <CLA>  
A:Cross-References: GB:M216; NID:9189729; PIDN:AA36427.1; PID:9189730  
R:Partanen, J.; Mäkelä, T.P.; Allitalo, R.; Lehteslahti, H.; Allitalo, K.  
Proc. Natl. Acad. Sci. U.S.A. 87, 8913-8917, 1990  
A:Title: Putative tyrosine kinases expressed in K-562 human leukemia cells.  
A:Reference number: A38268; MUID:91062389; PMID:2247464  
A:Accession: A38268  
A:Molecule type: mRNA  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 828-884 <PAR>  
R:Roberts, W.M.; Look, A.T.; Rousset, M.F.; Sherr, C.J.  
Cell 55, 655-661, 1988  
A:Title: Tandem linkage of human CSF-1 receptor (c-fms) and PDGF receptor genes.  
A:Reference number: A90908; MUID:89028677; PMID:2846185  
A:Accession: A31925  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 676-727 <ROB>  
A:Accession: B31925  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 901-932 <RO2>  
A:Accession: C31925  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1047-1106 <RO3>  
C:Comment: The extracellular domain is predicted to include five immunoglobulin-like dom  
C:Genetics:  
A:Gene: GDB:PDGFRB  
A:Cross-References: GDB:120710; OMIM:173410  
A:Map position: 5q31-q32  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
C:Keywords: ATP; autophosphorylation; glycoprotein; growth factor receptor; heterodimer;  
F:1-32/Domain: signal sequence #status predicted <SIG>  
F:33-1106/Product: platelet-derived growth factor receptor beta #status predicted <MAT>  
F:33-531/Domain: extracellular #status predicted <EXT>  
F:47-102/Domain: immunoglobulin homology <IM1>  
F:142-192/Domain: immunoglobulin homology <IM2>  
F:228-293/Domain: immunoglobulin homology <IM3>  
F:429-510/Domain: immunoglobulin homology <IM4>  
F:532-555/Domain: transmembrane #status predicted <TM>  
F:556-1106/Domain: intracellular #status predicted <INT>  
F:598-965/Domain: protein kinase homology <KIN>  
F:606-614/Region: protein kinase ATP-binding motif  
F:45-89,103,215,230,292,307,354,371,468,479/Binding site: carboxyrate (Asn) (covalent)  
F:54-100,149-190,235-291,436-508/Disulfide bonds: #status predicted  
F:634/Active site: Lys #status predicted  
F:857/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 40.0%; Score 6; DB 1; Length 1106;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGRH 6  
DB 656 SHLGRH 661

RESULT 41  
T01367  
hypothetical protein At2g34680 (imported) - Arabidopsis thaliana  
N:Alternate names: hypothetical protein T29F13.11  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 23-Mar-2001

C:Accession: T01367; F84759  
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; K  
submitted to the EMBL Data Library, May 1998  
A:Description: Arabidopsis thaliana chromosome II BAC T29F13 genomic sequence.  
A:Reference number: Z14179  
A:Accession: T01367  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1680 <ROU>  
A:Cross-References: EMBL:AC003096; NID:93132469; PID:93132477  
A:Experimental source: cultivar Columbia  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y  
M.; Ko, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umeyama, L.; Tallon,  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter  
Nature 402, 761-766, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: F84759  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1680 <STO>  
A:Cross-References: GB:AE002093; NID:93132477; PIDN:AA016266.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: T29F13.11; At2g34680  
A:Map position: 2  
A:Introns: 271/3; 295/3; 317/3; 341/3; 363/3; 384/3; 406/3; 440/1; 474/1; 551/3; 600/  
; 1372/2; 1433/1; 1485/1; 1529/1; 1548/3; 1578/3; 1622/3

Query Match 40.0%; Score 6; DB 2; Length 1680;  
Best Local Similarity 100.0%; Pred. No. 99;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPESR 13  
DB 248 STPESR 253

RESULT 42  
M47392  
chromodomain-helicase-DNA-binding protein, CHD-1 - mouse  
N:Alternate names: KYP protein  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A47392; S21568  
R:Delmas, V.; Stokes, D.G.; Perry, R.P.  
Proc. Natl. Acad. Sci. U.S.A. 90, 2414-2418, 1993  
A:Title: A mammalian DNA-binding protein that contains a chromodomain and an SNF2/SWI  
A:Reference number: A47392; MUID:93211972; PMID:8460153  
A:Accession: A47392  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1711 <DEU>  
A:Experimental source: S194 plasmacytoma cells  
A:Note: sequence inconsistent with the nucleotide translation  
A:Note: sequence extracted from NCBI backbone (NCBI:128272, NCBI:128273)  
R:Delmas, V.; Perry, R.P.  
submitted to the EMBL Data Library, May 1992  
A:Description: KYP, a mammalian protein that contains the SNF2/SWI2 helicase domain  
A:Reference number: S21568  
A:Accession: S21568  
A:Molecule type: mRNA  
A:Residues: 772-1711 <DE2>  
A:Cross-References: EMBL:X66028  
C:Superfamily: CHD-1 protein; chromobox homology  
C:Keywords: DNA binding  
F:293-336/Domain: chromobox homology <CB1>  
F:387-427/Domain: chromobox homology <CB2>

Query Match 40.0%; Score 6; DB 1; Length 1711;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HRSTPE 11

|||||  
Db 1698 RSTPE 1703

## RESULT 43

variant-specific surface protein 2 - malaria parasite (Plasmodium falciparum)  
C:Species: Plasmodium falciparum  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jun-2000  
C:Accession: T28626  
R:Su, X.Z.; Heatwole, V.M.; Wertheimer, S.P.; Guinet, F.; Herfeldt, J.A.; Peterson, D.S.  
Cell 82, 89-100, 1995  
A:Title: The large diverse gene family var encodes proteins involved in cytoadherence at  
A:Reference number: 220487; MUID:95330813; PMID:7606788  
A:Accession: T28626  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2664 <SUX>  
A:Cross-references: EMBL:L40609; NID:9886376; PID:9886378; PIDN:AAAT5398.1  
C:Genetics:  
A:Introns: 2197/3  
A>Note: var-2

Query Match 40.0%; Score 6; DB 2; Length 2664;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPES 12  
|||||  
Db 1335 RSTPES 1340

## RESULT 44

Insulin-like growth factor receptor type II - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 05-Nov-1999  
C:Accession: S58277  
R:Smrzka, O.W.; Stoger, R.; Kurzbauer, R.; Fae, I.; Fischer, G.F.; Barlow, D.P.  
submitted to the EMBL Data Library, January 1995  
A:Description: Conservation of a methylation imprint and a putative imprinting box at th  
A:Reference number: S58277  
A:Accession: S58277  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-18 <SMR>  
A:Cross-references: EMBL:X83702; NID:9929644; PIDN:CAAS8675.1; PID:9929645  
C:Keywords: growth factor receptor

Query Match 33.3%; Score 5; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15  
|||||  
Db 6 ESRAA 10

## RESULT 45

T04371  
thaumatin-like protein - barley (fragment)  
C:Species: Hordeum vulgare (barley)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 08-Oct-1999  
C:Accession: T04371  
R:Skadsen, R.W.; Herbst, J.M.  
submitted to the EMBL Data Library, July 1997  
A:Reference number: Z15316  
A:Accession: T04371  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-26 <SKA>  
A:Cross-references: EMBL:AF016328; NID:92454603; PIDN:AAAT1681.1; PID:92454604  
A:Experimental source: cv. Morex

C:Genetics:  
A:Gene: perm2

Query Match 33.3%; Score 5; DB 2; Length 26;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHR 7  
|||||  
Db 19 LGPHR 23

## RESULT 46

homeobox - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 15-Oct-1999  
C:Accession: J37554  
R:Moretti, P.; Simmons, P.; Thomas, P.; Haylock, D.; Rathjen, P.; Vadas, M.; D'Andrea  
Gene 144, 213-219, 1994  
A:Title: Identification of homeobox genes expressed in human haemopoietic progenitor  
A:Reference number: J37554; MUID:94314219; PMID:7518789  
A:Accession: J37554  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-39 <RES>  
A:Cross-references: EMBL:X74861; NID:9510992; PIDN:CAAS2854.1; PID:9510993  
C:Genetics:  
A:Gene: HPX-2  
C:Superfamily: unassigned homeobox proteins; homeobox homology  
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

Query Match 33.3%; Score 5; DB 2; Length 39;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13  
|||||  
Db 32 TPESR 36

## RESULT 47

F86491  
hypothetical protein CPJ0006 [imported] - Chlamydomonas pneumoniae (strain J138)  
C:Species: Chlamydomonas pneumoniae, Chlamydia pneumoniae  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001  
C:Accession: F86491  
R:Shiral, M.; Hiraoka, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;  
Nucleic Acids Res. 28, 2311-2314, 2000  
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
A:Reference number: A86491; MUID:20330349; PMID:10871362  
A:Accession: F86491  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-51 <STO>  
A:Cross-references: GB:BA000008; NID:98978379; PIDN:BAAG8216.1; GSPDB:GN00142  
A:Experimental source: strain J138  
C:Genetics:  
A:Gene: CPJ0006

Query Match 33.3%; Score 5; DB 2; Length 51;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11  
|||||  
Db 29 RSTPE 33

## RESULT 48

C72131  
hypothetical protein - Chlamydomonas pneumoniae (strain CWL029)  
C:Species: Chlamydomonas pneumoniae, Chlamydia pneumoniae



C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 05-May-2000  
 C:Accession: C72131  
 R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;  
 Nature Genet. 21, 385-389, 1999  
 A:Title: Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*.  
 A:Reference number: A72000; MUID:99206606; PMID:10192388  
 A:Accession: C72131  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-51 <ARN>  
 A:Cross-References: GB:AE001585; GB:AE001363; NID:g4376255; PIDN:AAD18164.1; PID:g437626  
 A:Experimental source: strain CWL029  
 C:Genetics:  
 A:Gene: Cpn0006

Query Match 33.3%; Score 5; DB 2; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 58;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11  
 Db 29 RSTPE 33

## RESULT 49

G72355

hypothetical protein - *Thermotoga maritima* (strain MSB8)C:Species: *Thermotoga maritima*

C&gt;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000

C:Accession: G72355

Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;  
 C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: G72355

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-56 &lt;ARN&gt;

A:Cross-References: GB:AE001735; GB:AE000512; NID:g4981122; PIDN:AAD35696.1; PID:g498113

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: TM0611

Query Match

Best Local Similarity 100.0%; Score 5; DB 2; Length 56;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESR 13

Db 22 TPESR 26

## RESULT 50

QOBCP7

ybeB protein - *Escherichia coli* (strain K-12)C:Species: *Escherichia coli*

C&gt;Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 01-Mar-2002

C:Accession: A24995; C64798

R:Asob, S.; Matsuzawa, H.; Ishino, F.; Strominger, J.L.; Matsubashi, M.; Ohta, T.

Eur. J. Biochem. 160, 231-238, 1986

A:Title: Nucleotide sequence of the ybeB gene and characteristics of the deduced amino a

A:Reference number: A91176; MUID:87030366; PMID:3533535

A:Accession: A24995

A:Molecule type: DNA

A:Residues: 1-69 &lt;ASO&gt;

A:Cross-References: GB:X04516; GB:D00001; GB:N00001; NID:g42313; PIDN:CAA28199.1; PID:g4

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of *Escherichia coli* K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: C64798

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-69 &lt;BIAT&gt;

A:Cross-References: GB:U00096; NID:g1786849; PIDN:MACT3738.1; PID:g17868

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: ybeB

A:Map position: 15 min

C:Superfamily: *Escherichia coli* ybeB protein

Query Match

Best Local Similarity 100.0%; Score 5; DB 1; Length 69;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRRA 15

Db 22 ESRRA 26

Search completed: February 6, 2003, 14:21:28  
 Job time : 19.9025 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 14:17:41 : Search time 26.1598 Seconds  
(Without alignments)  
525.254 Million cell updates/sec

Title: US-09-689-159A-2

Perfect score: 467

Sequence: 1 MTELPAPLSTYFQNAQMSDN.....ATDYLVQPFMDQLAFHQFYI 467

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size: 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database:

Issued\_patents\_AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep: \*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep: \*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep: \*  
5: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep: \*  
6: /cgn2\_6/ptodata/1/1aa/backfilltest.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	467	100.0	467	2	US-08-967-101-2
2	467	100.0	467	2	US-08-592-541-2
3	467	100.0	467	2	US-09-124-698-2
4	467	100.0	467	4	US-09-127-480-2
5	467	100.0	467	4	US-08-496-841C-2
6	467	100.0	467	4	US-09-124-523-2
7	438	93.8	463	3	US-08-888-077A-4
8	366	78.4	467	2	US-08-592-541-134
9	366	78.4	467	2	US-08-923-454A-10
10	366	78.4	467	3	US-08-670-964-2
11	366	78.4	467	3	US-08-888-077A-2
12	366	78.4	467	3	US-09-124-698-134
13	366	78.4	467	4	US-09-127-480-134
14	366	78.4	467	4	US-08-496-841C-134
15	366	78.4	467	4	US-08-832-867-3
16	366	78.4	467	4	US-09-227-725A-1
17	366	78.4	467	4	US-09-124-523-134
18	366	78.4	467	4	US-09-375-318-3
19	366	78.4	467	4	US-08-875-972-4
20	337	72.2	463	3	US-08-706-344C-2
21	306	65.5	407	2	US-08-706-344C-2
22	265	56.7	467	4	US-08-706-344C-2
23	204	43.7	467	4	US-08-706-344C-28
24	204	43.7	467	4	US-08-706-344C-30
25	204	43.7	467	4	US-08-706-344C-32
26	204	43.7	467	4	US-08-706-344C-32
27	189	40.5	462	3	US-08-788-231A-15

188	40.3	463	2	US-08-670-479-18	Sequence 18, Appl
29	20.8	467	2	US-08-967-101-4	Sequence 4, Appl
30	20.8	467	2	US-08-592-541-4	Sequence 17, Appl
31	20.8	467	3	US-08-888-077A-17	Sequence 4, Appl
32	20.8	467	3	US-09-124-698-4	Sequence 4, Appl
33	20.8	467	4	US-09-127-480-4	Sequence 4, Appl
34	20.8	467	4	US-08-496-841C-4	Sequence 4, Appl
35	20.8	467	4	US-08-496-841C-136	Sequence 136, Appl
36	20.8	467	4	US-09-124-523-4	Sequence 4, Appl
37	18.0	465	3	US-08-788-231A-17	Sequence 17, Appl
38	6.0	372	2	US-08-967-101-138	Sequence 138, Appl
39	6.0	372	2	US-08-592-541-138	Sequence 138, Appl
40	6.0	372	3	US-09-124-698-138	Sequence 138, Appl
41	6.0	372	4	US-09-127-480-138	Sequence 138, Appl
42	6.0	372	4	US-08-496-841C-138	Sequence 138, Appl
43	6.0	372	4	US-09-124-523-138	Sequence 138, Appl
44	6.0	447	2	US-08-875-972-29	Sequence 29, Appl
45	6.0	448	2	US-08-967-101-137	Sequence 137, Appl
46	6.0	448	2	US-08-592-541-137	Sequence 137, Appl
47	6.0	448	3	US-08-888-077A-19	Sequence 19, Appl
48	6.0	448	3	US-09-124-698-137	Sequence 137, Appl
49	6.0	448	4	US-09-127-480-137	Sequence 137, Appl
50	6.0	448	4	US-09-227-725A-2	Sequence 2, Appl
51	6.0	448	4	US-09-124-523-137	Sequence 137, Appl
52	6.0	448	4	US-09-375-318-2	Sequence 2, Appl
53	5.1	48	2	US-08-896-176-1	Sequence 1, Appl
54	4.9	29	2	US-08-896-176-4	Sequence 1, Appl
55	4.7	541	2	US-08-967-101-166	Sequence 166, Appl
56	4.7	541	2	US-08-592-541-166	Sequence 166, Appl
57	4.7	541	3	US-08-888-077A-21	Sequence 21, Appl
58	4.7	541	3	US-09-124-698-166	Sequence 166, Appl
59	4.7	541	4	US-09-127-480-166	Sequence 166, Appl
60	4.7	541	4	US-09-124-523-166	Sequence 166, Appl
61	4.3	21	2	US-08-896-176-2	Sequence 2, Appl
62	3.9	376	2	US-08-875-972-2	Sequence 2, Appl
63	3.2	15	2	US-08-967-101-167	Sequence 167, Appl
64	3.2	15	2	US-08-967-101-168	Sequence 168, Appl
65	3.2	15	2	US-08-967-101-169	Sequence 169, Appl
66	3.2	15	2	US-08-967-101-170	Sequence 170, Appl
67	3.2	15	2	US-08-592-541-167	Sequence 167, Appl
68	3.2	15	2	US-08-592-541-168	Sequence 168, Appl
69	3.2	15	2	US-08-592-541-169	Sequence 169, Appl
70	3.2	15	2	US-08-592-541-170	Sequence 170, Appl
71	3.2	15	3	US-09-124-698-167	Sequence 167, Appl
72	3.2	15	3	US-09-124-698-168	Sequence 168, Appl
73	3.2	15	3	US-09-124-698-169	Sequence 169, Appl
74	3.2	15	3	US-09-124-698-170	Sequence 170, Appl
75	3.2	15	4	US-08-496-841C-164	Sequence 164, Appl
76	3.2	15	4	US-08-496-841C-165	Sequence 165, Appl
77	3.2	15	4	US-08-496-841C-166	Sequence 166, Appl
78	3.2	15	4	US-08-496-841C-167	Sequence 167, Appl
79	3.2	15	4	US-09-124-523-167	Sequence 167, Appl
80	3.2	15	4	US-09-124-523-168	Sequence 168, Appl
81	3.2	15	4	US-09-124-523-169	Sequence 169, Appl
82	3.2	15	4	US-09-124-523-170	Sequence 170, Appl
83	3.2	15	4	US-08-896-176-3	Sequence 3, Appl
84	3.2	20	4	US-09-375-318-41	Sequence 41, Appl
85	3.2	20	4	US-08-896-176-9	Sequence 9, Appl
86	3.2	15	4	US-08-784-649A-2	Sequence 2, Appl
87	3.2	1280	2	US-08-583-276-19	Sequence 19, Appl
88	3.2	1280	2	US-08-752-447-2	Sequence 2, Appl
89	3.2	1280	4	US-09-316-167-2	Sequence 2, Appl
90	3.2	1280	6	5206352-4	Patent No. 5206352
91	1.9	1279	2	US-08-090-036-1	Sequence 1, Appl
92	1.7	1280	2	US-09-319-730-17	Sequence 17, Appl
93	1.7	1280	4	US-09-319-730-15	Sequence 15, Appl
94	1.7	1280	4	US-09-319-730-19	Sequence 19, Appl
95	1.7	1280	4	US-09-177-249-56	Sequence 56, Appl
96	1.5	31	1		
97	1.5	38	4		
98	1.5	61	4		
99	1.5	74	4		
100	1.5	74	4		

101	7	1.5	118	2	US-08-767-026-4	Sequence 4, Appli	174	6	1.3	12	5	PCT-US95-02044-22	Sequence 22, Appl
102	7	1.5	118	2	US-08-846-021A-2	Sequence 2, Appli	175	6	1.3	14	3	US-08-199-776-18	Sequence 18, Appl
103	7	1.5	118	2	US-08-846-021A-7	Sequence 7, Appli	176	6	1.3	14	1	US-08-663-731-18	Sequence 18, Appl
104	7	1.5	154	1	US-08-366-783-5	Sequence 5, Appli	177	6	1.3	14	3	US-08-879-338-18	Sequence 18, Appl
105	7	1.5	173	1	US-08-366-783-2	Sequence 2, Appli	178	6	1.3	14	4	US-09-293-238B-18	Sequence 18, Appl
106	7	1.5	173	1	US-08-313-098A-2	Sequence 2, Appli	179	6	1.3	14	5	PCT-US95-02044-18	Sequence 18, Appl
107	7	1.5	173	2	US-08-767-026-2	Sequence 2, Appli	180	6	1.3	15	1	US-08-199-776-20	Sequence 20, Appl
108	7	1.5	173	2	US-08-831-575-6	Sequence 6, Appli	181	6	1.3	15	3	US-08-663-731-20	Sequence 20, Appl
109	7	1.5	187	2	US-08-846-021A-5	Sequence 5, Appli	182	6	1.3	15	3	US-08-879-338-20	Sequence 20, Appl
110	7	1.5	202	4	US-09-615-192A-386	Sequence 386, App	183	6	1.3	15	4	US-09-009-953-234	Sequence 234, App
111	7	1.5	213	4	US-09-134-001C-3556	Sequence 3556, Ap	184	6	1.3	15	4	US-09-009-953-252	Sequence 252, App
112	7	1.5	231	2	US-08-808-550-33	Sequence 33, Appli	185	6	1.3	15	4	US-09-293-238B-20	Sequence 20, Appl
113	7	1.5	234	1	US-08-366-783-4	Sequence 4, Appli	186	6	1.3	15	5	PCT-US95-02044-20	Sequence 20, Appl
114	7	1.5	254	2	US-08-767-026-7	Sequence 7, Appli	187	6	1.3	16	4	US-08-602-999A-220	Sequence 220, App
115	7	1.5	323	4	US-09-292-858B-22	Sequence 22, Appli	188	6	1.3	16	4	US-09-500-124-220	Sequence 220, App
116	7	1.5	366	4	US-09-210-843-2	Sequence 2, Appli	189	6	1.3	19	4	US-09-177-249-232	Sequence 232, App
117	7	1.5	392	4	US-09-026-408-2	Sequence 2, Appli	190	6	1.3	20	1	US-08-199-776-21	Sequence 21, Appl
118	7	1.5	405	4	US-09-026-408-13	Sequence 13, Appli	191	6	1.3	20	3	US-08-663-731-21	Sequence 21, Appl
119	7	1.5	434	2	US-08-989-925-1	Sequence 1, Appli	192	6	1.3	20	3	US-08-879-338-21	Sequence 21, Appl
120	7	1.5	482	3	US-09-135-639-2	Sequence 2, Appli	193	6	1.3	21	4	US-09-293-238B-21	Sequence 21, Appl
121	7	1.5	527	4	US-09-370-838-216	Sequence 216, App	194	6	1.3	20	5	PCT-US95-02044-21	Sequence 21, Appl
122	7	1.5	564	4	US-09-134-001C-5139	Sequence 5139, Ap	195	6	1.3	21	1	US-08-199-776-5	Sequence 5, Appli
123	7	1.5	571	4	US-09-134-001C-3865	Sequence 3865, Ap	196	6	1.3	21	3	US-08-663-731-5	Sequence 5, Appli
124	7	1.5	689	4	US-09-177-249-2	Sequence 2, Appli	197	6	1.3	21	3	US-08-879-338-5	Sequence 5, Appli
125	7	1.5	689	4	US-09-061-769A-2	Sequence 2, Appli	198	6	1.3	21	3	US-08-879-338-26	Sequence 26, Appli
126	7	1.5	714	4	US-08-990-114-3	Sequence 3, Appli	199	6	1.3	21	4	US-09-336-536-25	Sequence 25, Appli
127	7	1.5	714	4	US-09-241-333-3	Sequence 3, Appli	200	6	1.3	21	4	US-09-293-238B-5	Sequence 5, Appli
128	7	1.5	740	4	US-09-022-983-5	Sequence 5, Appli	201	6	1.3	21	4	US-09-293-238B-26	Sequence 26, Appli
129	7	1.5	830	4	US-09-562-737-31	Sequence 31, Appli	202	6	1.3	21	5	PCT-US95-02044-5	Sequence 5, Appli
130	7	1.5	830	4	US-09-562-737-33	Sequence 33, Appli	203	6	1.3	25	4	US-09-149-876-431	Sequence 431, App
131	7	1.5	830	4	US-09-562-737-34	Sequence 34, Appli	204	6	1.3	25	4	US-09-911-927-35	Sequence 35, Appli
132	7	1.5	830	4	US-09-562-737-35	Sequence 35, Appli	205	6	1.3	25	4	US-09-911-882-35	Sequence 35, Appli
133	7	1.5	830	4	US-09-562-737-36	Sequence 36, Appli	206	6	1.3	26	4	US-08-965-762-35	Sequence 35, Appli
134	7	1.5	979	4	US-08-878-474-5	Sequence 5, Appli	207	6	1.3	27	4	US-09-227-357-642	Sequence 642, App
135	7	1.5	1023	2	US-08-475-891A-2	Sequence 2, Appli	208	6	1.3	29	4	US-09-143-124-24	Sequence 24, Appli
136	7	1.5	1023	2	US-08-567-375-2	Sequence 2, Appli	209	6	1.3	32	1	US-08-190-802A-119	Sequence 119, App
137	7	1.5	1023	2	US-08-587-600A-2	Sequence 2, Appli	210	6	1.3	32	4	US-08-477-946-119	Sequence 119, App
138	7	1.5	1162	2	US-08-728-333A-2	Sequence 2, Appli	211	6	1.3	32	4	US-08-473-089-119	Sequence 119, App
139	7	1.5	1162	2	US-09-298-568-2	Sequence 2, Appli	212	6	1.3	32	4	US-08-487-072A-119	Sequence 119, App
140	7	1.5	2647	2	US-08-583-562B-8	Sequence 8, Appli	213	6	1.3	37	4	US-09-149-476-694	Sequence 694, App
141	7	1.5	2647	2	US-08-779-113-8	Sequence 8, Appli	214	6	1.3	39	5	PCT-US92-07813-9	Sequence 9, Appli
142	7	1.5	2987	2	US-08-970-269A-29	Sequence 29, Appli	215	6	1.3	41	2	US-08-366-800-3	Sequence 3, Appli
143	7	1.5	2987	2	US-09-407-562-29	Sequence 29, Appli	216	6	1.3	43	2	US-08-499-676A-29	Sequence 29, Appli
144	7	1.5	3959	2	US-08-970-269A-30	Sequence 30, Appli	217	6	1.3	43	2	US-08-499-676A-30	Sequence 30, Appli
145	7	1.5	3959	2	US-09-407-562-30	Sequence 30, Appli	218	6	1.3	43	2	US-08-499-676A-31	Sequence 31, Appli
146	6	1.3	6	3	US-08-738-381-49	Sequence 49, Appli	219	6	1.3	43	4	US-09-314-268-157	Sequence 157, App
147	6	1.3	6	3	US-08-738-381-50	Sequence 50, Appli	220	6	1.3	44	4	US-08-905-223-345	Sequence 345, App
148	6	1.3	6	3	US-08-738-381-52	Sequence 52, Appli	221	6	1.3	45	1	US-08-062-472B-24	Sequence 24, Appli
149	6	1.3	6	1	US-08-199-776-13	Sequence 13, Appli	222	6	1.3	46	4	US-09-227-357-278	Sequence 278, App
150	6	1.3	10	1	US-08-199-776-14	Sequence 14, Appli	223	6	1.3	50	4	US-09-314-268-160	Sequence 160, App
151	6	1.3	10	1	US-08-199-776-15	Sequence 15, Appli	224	6	1.3	55	1	US-08-199-776-4	Sequence 4, Appli
152	6	1.3	10	1	US-08-199-776-16	Sequence 16, Appli	225	6	1.3	55	3	US-08-663-731-4	Sequence 4, Appli
153	6	1.3	10	3	US-08-663-731-13	Sequence 13, Appli	226	6	1.3	55	3	US-08-879-338-4	Sequence 4, Appli
154	6	1.3	10	3	US-08-663-731-14	Sequence 14, Appli	227	6	1.3	55	4	US-09-293-238B-4	Sequence 4, Appli
155	6	1.3	10	3	US-08-663-731-15	Sequence 15, Appli	228	6	1.3	55	5	PCT-US95-02044-4	Sequence 4, Appli
156	6	1.3	10	3	US-08-663-731-16	Sequence 16, Appli	229	6	1.3	60	3	US-08-955-937A-6	Sequence 6, Appli
157	6	1.3	10	3	US-08-879-338-13	Sequence 13, Appli	230	6	1.3	62	1	US-09-300-985-6	Sequence 6, Appli
158	6	1.3	10	3	US-08-879-338-14	Sequence 14, Appli	231	6	1.3	62	1	US-08-083-947-4	Sequence 4, Appli
159	6	1.3	10	3	US-08-879-338-15	Sequence 15, Appli	232	6	1.3	62	1	US-08-083-947-9	Sequence 9, Appli
160	6	1.3	10	3	US-08-879-338-16	Sequence 16, Appli	233	6	1.3	62	1	US-08-083-947-13	Sequence 13, Appli
161	6	1.3	10	4	US-09-293-238B-13	Sequence 13, Appli	234	6	1.3	62	5	PCT-US94-07088-4	Sequence 4, Appli
162	6	1.3	10	4	US-09-293-238B-14	Sequence 14, Appli	235	6	1.3	62	5	PCT-US94-07088-9	Sequence 9, Appli
163	6	1.3	10	4	US-09-293-238B-15	Sequence 15, Appli	236	6	1.3	62	5	PCT-US94-07088-13	Sequence 13, Appli
164	6	1.3	10	4	US-09-293-238B-16	Sequence 16, Appli	237	6	1.3	71	1	US-08-083-947-2	Sequence 2, Appli
165	6	1.3	10	5	PCT-US95-02044-13	Sequence 13, Appli	238	6	1.3	71	1	US-08-083-947-5	Sequence 5, Appli
166	6	1.3	10	5	PCT-US95-02044-14	Sequence 14, Appli	239	6	1.3	71	1	US-08-083-947-10	Sequence 10, Appli
167	6	1.3	10	5	PCT-US95-02044-15	Sequence 15, Appli	240	6	1.3	71	1	US-08-083-947-14	Sequence 14, Appli
168	6	1.3	10	5	PCT-US95-02044-16	Sequence 16, Appli	241	6	1.3	71	5	PCT-US94-07088-2	Sequence 2, Appli
169	6	1.3	12	3	US-08-199-776-22	Sequence 22, Appli	242	6	1.3	71	5	PCT-US94-07088-5	Sequence 5, Appli
170	6	1.3	12	3	US-08-663-731-22	Sequence 22, Appli	243	6	1.3	71	5	PCT-US94-07088-10	Sequence 10, Appli
171	6	1.3	12	3	US-08-879-338-22	Sequence 22, Appli	244	6	1.3	71	5	PCT-US94-07088-14	Sequence 14, Appli
172	6	1.3	12	4	US-09-752-165-53	Sequence 53, Appli	245	6	1.3	77	1	US-08-530-550-6	Sequence 6, Appli
173	6	1.3	12	4	US-09-293-238B-22	Sequence 22, Appli	246	6	1.3	77	5	PCT-US95-13660-6	Sequence 6, Appli

247	6	1.3	77	6	5175095-5	Patent No. 5175095	320	6	1.3	137	4	US-09-325-932A-170	Sequence 170, App
248	6	1.3	80	4	US-09-145-828A-3	Sequence 3, Appl1	321	6	1.3	140	2	US-08-483-636-4	Sequence 4, Appl1
249	6	1.3	81	1	US-08-083-947-1	Sequence 1, Appl1	322	6	1.3	140	2	US-08-483-637-4	Sequence 4, Appl1
250	6	1.3	81	1	US-08-083-947-6	Sequence 6, Appl1	323	6	1.3	142	4	US-09-134-001C-4136	Sequence 4136, Ap
251	6	1.3	81	1	US-08-083-947-11	Sequence 11, Appl1	324	6	1.3	144	2	US-08-186-895-10	Sequence 10, Appl1
252	6	1.3	81	1	US-08-083-947-15	Sequence 15, Appl1	325	6	1.3	144	2	US-08-888-497-37	Sequence 37, Appl1
253	6	1.3	81	1	US-08-530-550-4	Sequence 4, Appl1	326	6	1.3	144	4	US-09-362-230-37	Sequence 37, Appl1
254	6	1.3	81	1	US-08-530-550-7	Sequence 7, Appl1	327	6	1.3	144	5	PCR-US94-07926-37	Sequence 37, Appl1
255	6	1.3	81	1	US-08-530-550-8	Sequence 8, Appl1	328	6	1.3	144	6	5187153-18	Patent No. 5187153
256	6	1.3	81	1	US-08-530-550-9	Sequence 9, Appl1	329	6	1.3	145	2	US-08-832-535-11	Sequence 11, Appl1
257	6	1.3	81	1	US-08-530-550-10	Sequence 10, Appl1	330	6	1.3	146	2	US-08-791-522-3	Sequence 3, Appl1
258	6	1.3	81	1	US-08-530-550-11	Sequence 11, Appl1	331	6	1.3	146	3	US-08-744-138-3	Sequence 4, Appl1
259	6	1.3	81	1	US-08-530-550-12	Sequence 12, Appl1	332	6	1.3	146	3	US-09-019-485-4	Sequence 4, Appl1
260	6	1.3	81	1	US-08-530-550-13	Sequence 13, Appl1	333	6	1.3	146	3	US-09-314-777-3	Sequence 3, Appl1
261	6	1.3	81	1	US-08-530-550-14	Sequence 14, Appl1	334	6	1.3	146	4	US-09-431-480-6	Sequence 6, Appl1
262	6	1.3	81	1	US-08-530-550-15	Sequence 15, Appl1	335	6	1.3	146	4	US-09-617-302-6	Sequence 6, Appl1
263	6	1.3	81	5	PCR-US94-07088-1	Sequence 1, Appl1	336	6	1.3	146	4	US-09-241-376-3	Sequence 3, Appl1
264	6	1.3	81	5	PCR-US94-07088-6	Sequence 6, Appl1	337	6	1.3	146	5	PCR-US95-07135-9	Sequence 9, Appl1
265	6	1.3	81	5	PCR-US94-07088-11	Sequence 11, Appl1	338	6	1.3	146	6	5432264-6	Patent No. 5432264
266	6	1.3	81	5	PCR-US95-13660-15	Sequence 15, Appl1	339	6	1.3	151	2	US-08-637-759B-349	Sequence 349, App
267	6	1.3	81	5	PCR-US95-13660-4	Sequence 4, Appl1	340	6	1.3	151	3	US-08-871-355A-349	Sequence 349, App
268	6	1.3	81	5	PCR-US95-13660-7	Sequence 7, Appl1	341	6	1.3	151	4	US-09-201-942-349	Sequence 349, App
269	6	1.3	81	5	PCR-US95-13660-8	Sequence 8, Appl1	342	6	1.3	153	3	US-08-943-173-1	Sequence 1, Appl1
270	6	1.3	81	5	PCR-US95-13660-9	Sequence 9, Appl1	343	6	1.3	155	1	US-08-530-010-13	Sequence 13, Appl1
271	6	1.3	81	5	PCR-US95-13660-10	Sequence 10, Appl1	344	6	1.3	155	2	US-08-484-101B-13	Sequence 13, Appl1
272	6	1.3	81	5	PCR-US95-13660-11	Sequence 11, Appl1	345	6	1.3	155	4	US-08-714-524D-13	Sequence 13, Appl1
273	6	1.3	81	5	PCR-US95-13660-12	Sequence 12, Appl1	346	6	1.3	157	2	US-08-883-070-3	Sequence 3, Appl1
274	6	1.3	81	5	PCR-US95-13660-13	Sequence 13, Appl1	347	6	1.3	159	1	US-07-853-985A-8	Sequence 8, Appl1
275	6	1.3	81	5	PCR-US95-13660-14	Sequence 14, Appl1	348	6	1.3	159	1	US-07-681-703B-8	Sequence 8, Appl1
276	6	1.3	81	5	PCR-US95-13660-15	Sequence 15, Appl1	349	6	1.3	159	1	US-08-184-236-8	Sequence 8, Appl1
277	6	1.3	83	2	US-08-499-676A-10	Sequence 10, Appl1	350	6	1.3	159	2	US-08-407-410B-8	Sequence 8, Appl1
278	6	1.3	83	2	US-08-499-676A-11	Sequence 11, Appl1	351	6	1.3	159	2	US-08-485-500-8	Sequence 8, Appl1
279	6	1.3	83	2	US-08-499-676A-12	Sequence 12, Appl1	352	6	1.3	159	5	PCR-US91-02370-8	Sequence 8, Appl1
280	6	1.3	84	1	US-08-083-947-8	Sequence 8, Appl1	353	6	1.3	159	5	PCR-US94-04174-8	Sequence 8, Appl1
281	6	1.3	84	1	US-08-530-550-21	Sequence 21, Appl1	354	6	1.3	160	4	US-09-247-155-156	Sequence 156, App
282	6	1.3	84	5	PCR-US94-07088-8	Sequence 8, Appl1	355	6	1.3	162	1	US-07-879-685B-4	Sequence 4, Appl1
283	6	1.3	84	5	PCR-US95-13660-21	Sequence 21, Appl1	356	6	1.3	170	2	US-08-820-170A-22	Sequence 22, Appl1
284	6	1.3	89	4	US-08-444-818-22	Sequence 4, Appl1	357	6	1.3	170	2	US-08-918-723-1	Sequence 1, Appl1
285	6	1.3	92	4	US-08-905-223-408	Sequence 408, App	358	6	1.3	170	2	US-09-237-507-1	Sequence 1, Appl1
286	6	1.3	92	4	US-09-394-630-2	Sequence 2, Appl1	359	6	1.3	170	3	US-09-055-699-22	Sequence 22, Appl1
287	6	1.3	93	1	US-08-083-947-7	Sequence 7, Appl1	360	6	1.3	170	3	US-09-069-896-1	Sequence 1, Appl1
288	6	1.3	93	1	US-08-083-947-12	Sequence 12, Appl1	361	6	1.3	170	4	US-09-273-565-22	Sequence 22, Appl1
289	6	1.3	93	5	PCR-US94-07088-7	Sequence 7, Appl1	362	6	1.3	170	4	US-09-565-538-22	Sequence 22, Appl1
290	6	1.3	93	5	PCR-US94-07088-12	Sequence 12, Appl1	363	6	1.3	170	4	US-09-661-468-22	Sequence 22, Appl1
291	6	1.3	98	4	US-09-020-846-40	Sequence 40, Appl1	364	6	1.3	170	4	US-09-471-468-1	Sequence 1, Appl1
292	6	1.3	102	1	US-08-350-884-80	Sequence 80, Appl1	365	6	1.3	173	2	US-08-537-811-43	Sequence 43, Appl1
293	6	1.3	102	1	US-08-709-173-80	Sequence 80, Appl1	366	6	1.3	178	2	US-08-487-031-23	Sequence 23, Appl1
294	6	1.3	102	2	US-08-411-913-9	Sequence 9, Appl1	367	6	1.3	178	3	US-08-473-034-23	Sequence 23, Appl1
295	6	1.3	102	2	US-08-709-177-80	Sequence 80, Appl1	368	6	1.3	178	4	US-08-259-451-15	Sequence 15, Appl1
296	6	1.3	102	4	US-08-444-818-144	Sequence 144, App	369	6	1.3	180	1	US-08-062-472B-7	Sequence 7, Appl1
297	6	1.3	105	2	US-08-487-031-20	Sequence 20, Appl1	370	6	1.3	182	1	US-08-466-603-5	Sequence 5, Appl1
298	6	1.3	105	3	US-08-473-034-20	Sequence 20, Appl1	371	6	1.3	182	1	US-08-314-503A-5	Sequence 5, Appl1
299	6	1.3	105	4	US-09-149-476-479	Sequence 479, App	372	6	1.3	182	1	US-08-468-066-5	Sequence 5, Appl1
300	6	1.3	106	1	US-08-290-919-11	Sequence 11, Appl1	373	6	1.3	182	2	US-08-466-717-5	Sequence 5, Appl1
301	6	1.3	107	2	US-08-647-144-6	Sequence 6, Appl1	374	6	1.3	182	3	PCR-US95-12414-5	Sequence 5, Appl1
302	6	1.3	115	4	US-09-145-828A-15	Sequence 15, Appl1	375	6	1.3	182	5	PCR-US95-12414-5	Sequence 5, Appl1
303	6	1.3	116	3	US-08-621-018B-4	Sequence 4, Appl1	376	6	1.3	186	1	US-07-853-985A-10	Sequence 10, Appl1
304	6	1.3	120	4	US-08-751-359-25	Sequence 25, Appl1	377	6	1.3	186	1	US-07-681-703B-10	Sequence 10, Appl1
305	6	1.3	120	4	US-08-907-146-25	Sequence 25, Appl1	378	6	1.3	186	1	US-08-184-236-10	Sequence 10, Appl1
306	6	1.3	123	1	US-08-398-613A-20	Sequence 20, Appl1	379	6	1.3	186	2	US-08-407-410B-10	Sequence 10, Appl1
307	6	1.3	123	1	US-08-398-612A-20	Sequence 20, Appl1	380	6	1.3	186	2	US-08-485-500-10	Sequence 10, Appl1
308	6	1.3	123	1	US-08-398-611A-20	Sequence 20, Appl1	381	6	1.3	186	5	PCR-US91-02370-10	Sequence 10, Appl1
309	6	1.3	123	2	US-08-491-334A-20	Sequence 20, Appl1	382	6	1.3	186	5	PCR-US94-04174-10	Sequence 10, Appl1
310	6	1.3	123	2	US-08-822-573-2	Sequence 2, Appl1	383	6	1.3	190	4	US-08-918-428D-3	Sequence 3, Appl1
311	6	1.3	123	3	US-09-027-449-17	Sequence 17, Appl1	384	6	1.3	194	4	US-08-959-004-9	Sequence 9, Appl1
312	6	1.3	123	3	US-08-804-444A-17	Sequence 17, Appl1	385	6	1.3	196	3	US-09-122-443-4	Sequence 4, Appl1
313	6	1.3	123	4	US-09-026-985-17	Sequence 17, Appl1	386	6	1.3	199	4	US-09-134-001C-4391	Sequence 4391, Ap
314	6	1.3	123	4	US-09-121-952A-17	Sequence 17, Appl1	387	6	1.3	200	4	US-08-469-260A-395	Sequence 395, App
315	6	1.3	123	4	US-09-234-340A-17	Sequence 17, Appl1	388	6	1.3	201	2	US-08-716-317-1	Sequence 1, Appl1
316	6	1.3	123	6	5466783-22	Patent No. 5466783	389	6	1.3	207	3	US-08-559-397A-13	Sequence 13, Appl1
317	6	1.3	124	4	US-08-307-499-2	Sequence 2, Appl1	390	6	1.3	209	3	US-08-559-397A-11	Sequence 11, Appl1
318	6	1.3	124	4	US-09-299-268-2	Sequence 2, Appl1	391	6	1.3	212	4	US-08-937-067-4	Sequence 4, Appl1
319	6	1.3	125	4	US-08-905-223-320	Sequence 320, App	392	6	1.3	219	1	US-08-441-629-13	Sequence 13, Appl1

393	6	1.3	219	3	US-08-776-207-13	Sequence 13, Appl	466	6	1.3	253	3	US-09-146-249A-46	Sequence 46, Appl
394	6	1.3	219	4	US-09-507-773-13	Sequence 13, Appl	467	6	1.3	253	3	US-08-206-188B-46	Sequence 46, Appl
395	6	1.3	219	5	PCT-US95-09172-13	Sequence 13, Appl	468	6	1.3	254	4	US-09-134-001C-4582	Sequence 4582, Ap
396	6	1.3	222	1	US-07-688-352C-46	Sequence 46, Appl	469	6	1.3	256	4	US-09-325-932A-57	Sequence 57, Appl
397	6	1.3	222	4	US-09-134-001C-4748	Sequence 4748, Ap	470	6	1.3	260	4	US-09-216-295-23	Sequence 23, Appl
398	6	1.3	222	5	PCT-US91-022714-43	Sequence 43, Appl	471	6	1.3	263	4	US-09-210-295-24	Sequence 24, Appl
399	6	1.3	222	5	PCT-US91-022714-54	Sequence 54, Appl	472	6	1.3	266	4	US-08-444-818-32	Sequence 32, Appl
400	6	1.3	228	1	US-08-278-091-10	Sequence 10, Appl	473	6	1.3	266	4	US-08-850-328-7	Sequence 7, Appl
401	6	1.3	228	1	US-08-483-859-10	Sequence 10, Appl	474	6	1.3	266	4	US-09-417-721-9	Sequence 9, Appl
402	6	1.3	228	1	US-08-472-173-10	Sequence 10, Appl	475	6	1.3	266	4	US-09-417-721-15	Sequence 15, Appl
403	6	1.3	228	2	US-08-487-167-10	Sequence 10, Appl	476	6	1.3	266	6	5175383-5	Patent No. 5175383
404	6	1.3	228	2	US-08-482-816-10	Sequence 10, Appl	477	6	1.3	267	1	US-08-462-159B-13	Sequence 13, Appl
405	6	1.3	228	2	US-08-296-149-10	Sequence 10, Appl	478	6	1.3	267	3	US-09-103-079-13	Sequence 13, Appl
406	6	1.3	228	2	US-08-801-499-10	Sequence 10, Appl	479	6	1.3	267	4	US-08-871-572B-7	Sequence 7, Appl
407	6	1.3	228	2	US-08-615-271-10	Sequence 10, Appl	480	6	1.3	268	1	US-08-439-725A-12	Sequence 12, Appl
408	6	1.3	228	3	US-09-074-660-10	Sequence 10, Appl	481	6	1.3	268	1	US-08-464-590A-17	Sequence 17, Appl
409	6	1.3	228	3	US-09-074-659-10	Sequence 10, Appl	482	6	1.3	268	2	US-08-207-412B-12	Sequence 12, Appl
410	6	1.3	228	3	US-09-106-468-10	Sequence 10, Appl	483	6	1.3	268	2	US-08-867-471-12	Sequence 12, Appl
411	6	1.3	228	4	US-09-106-468A-10	Sequence 10, Appl	484	6	1.3	268	2	US-08-438-439C-8	Sequence 8, Appl
412	6	1.3	228	4	US-09-106-467-10	Sequence 10, Appl	485	6	1.3	268	2	US-08-951-822-33	Sequence 33, Appl
413	6	1.3	231	1	US-08-220-379B-7	Sequence 7, Appl	486	6	1.3	268	3	US-08-705-245-14	Sequence 14, Appl
414	6	1.3	231	1	US-08-243-545-2	Sequence 2, Appl	487	6	1.3	268	3	US-08-718-904-14	Sequence 14, Appl
415	6	1.3	231	2	US-08-993-962-2	Sequence 2, Appl	488	6	1.3	268	3	US-09-023-082A-15	Sequence 15, Appl
416	6	1.3	231	4	US-09-160-841-2	Sequence 2, Appl	489	6	1.3	268	3	US-09-093-585-17	Sequence 17, Appl
417	6	1.3	231	5	PCT-US94-05365-2	Sequence 2, Appl	490	6	1.3	268	4	US-09-240-952-2	Sequence 2, Appl
418	6	1.3	231	5	PCT-US95-03866-6	Sequence 6, Appl	491	6	1.3	268	4	US-09-368-951-13	Sequence 33, Appl
419	6	1.3	233	1	US-08-032-848C-12	Sequence 12, Appl	492	6	1.3	269	2	US-08-438-439C-18	Sequence 18, Appl
420	6	1.3	235	1	US-08-836-236-6	Sequence 6, Appl	493	6	1.3	270	2	US-07-857-224B-76	Sequence 76, Appl
421	6	1.3	236	1	US-08-118-270-74	Sequence 74, Appl	494	6	1.3	270	4	US-09-399-913-59	Sequence 59, Appl
422	6	1.3	236	5	PCT-US93-08528-74	Sequence 74, Appl	495	6	1.3	272	1	US-08-350-884-84	Sequence 84, Appl
423	6	1.3	237	1	US-08-398-612A-28	Sequence 28, Appl	496	6	1.3	272	1	US-08-709-173-84	Sequence 84, Appl
424	6	1.3	237	1	US-08-398-611A-28	Sequence 28, Appl	497	6	1.3	272	2	US-08-709-177-84	Sequence 84, Appl
425	6	1.3	237	2	US-08-491-334A-28	Sequence 28, Appl	498	6	1.3	280	4	US-09-720-817-1	Sequence 1, Appl
426	6	1.3	237	3	US-09-027-449-25	Sequence 25, Appl	499	6	1.3	280	4	US-09-145-828A-21	Sequence 21, Appl
427	6	1.3	237	3	US-08-804-444A-25	Sequence 25, Appl	500	6	1.3	283	4	US-09-134-001C-5491	Sequence 5491, Ap
428	6	1.3	237	4	US-09-026-985-25	Sequence 25, Appl	501	6	1.3	289	4	US-09-145-828A-17	Sequence 17, Appl
429	6	1.3	237	4	US-09-121-952A-25	Sequence 25, Appl	502	6	1.3	293	4	US-09-145-828A-12	Sequence 12, Appl
430	6	1.3	237	4	US-09-234-340A-25	Sequence 25, Appl	503	6	1.3	295	3	US-08-511-759B-2	Sequence 2, Appl
431	6	1.3	238	1	US-08-398-613A-28	Sequence 28, Appl	504	6	1.3	295	4	US-09-592-197-2	Sequence 2, Appl
432	6	1.3	238	4	US-09-111-470-8	Sequence 8, Appl	505	6	1.3	295	4	US-08-892-704-2	Sequence 2, Appl
433	6	1.3	240	1	US-08-278-091-11	Sequence 11, Appl	506	6	1.3	296	4	US-09-134-001C-4473	Sequence 4473, Ap
434	6	1.3	240	1	US-08-483-859-11	Sequence 11, Appl	507	6	1.3	297	3	US-09-187-049-10	Sequence 10, Appl
435	6	1.3	240	1	US-08-472-173-11	Sequence 11, Appl	508	6	1.3	297	3	US-09-134-001C-3397	Sequence 3397, Ap
436	6	1.3	240	2	US-08-114-555A-6	Sequence 6, Appl	509	6	1.3	298	1	US-08-393-985-10	Sequence 10, Appl
437	6	1.3	240	2	US-08-114-555A-8	Sequence 8, Appl	510	6	1.3	298	4	US-09-134-001C-3631	Sequence 3631, Ap
438	6	1.3	240	2	US-08-487-167-11	Sequence 11, Appl	511	6	1.3	300	4	US-09-194-146-6	Sequence 6, Appl
439	6	1.3	240	2	US-08-482-816-11	Sequence 11, Appl	512	6	1.3	302	3	US-08-511-759B-9	Sequence 9, Appl
440	6	1.3	240	2	US-08-296-149-11	Sequence 11, Appl	513	6	1.3	302	4	US-09-592-197-9	Sequence 9, Appl
441	6	1.3	240	2	US-08-801-499-11	Sequence 11, Appl	514	6	1.3	302	4	US-08-892-704-9	Sequence 9, Appl
442	6	1.3	240	2	US-08-760-745-3	Sequence 3, Appl	515	6	1.3	313	3	US-08-926-842B-62	Sequence 62, Appl
443	6	1.3	240	2	US-08-615-271-11	Sequence 11, Appl	516	6	1.3	315	4	US-09-720-817-3	Sequence 3, Appl
444	6	1.3	240	3	US-09-074-660-11	Sequence 11, Appl	517	6	1.3	315	4	US-09-720-817-6	Sequence 6, Appl
445	6	1.3	240	3	US-09-074-659-11	Sequence 11, Appl	518	6	1.3	317	4	US-09-145-828A-7	Sequence 7, Appl
446	6	1.3	240	3	US-08-559-397A-12	Sequence 12, Appl	519	6	1.3	318	4	US-08-878-474-3	Sequence 3, Appl
447	6	1.3	240	3	US-08-559-397A-14	Sequence 14, Appl	520	6	1.3	318	4	US-09-483-371-2	Sequence 2, Appl
448	6	1.3	240	3	US-09-106-468-11	Sequence 11, Appl	521	6	1.3	318	4	US-09-145-828A-19	Sequence 19, Appl
449	6	1.3	240	4	US-09-106-466A-11	Sequence 11, Appl	522	6	1.3	319	4	US-09-008-697A-6	Sequence 6, Appl
450	6	1.3	240	4	US-09-106-467-11	Sequence 11, Appl	523	6	1.3	320	1	US-08-365-103B-10	Sequence 10, Appl
451	6	1.3	246	4	US-08-634-475-3	Sequence 3, Appl	524	6	1.3	320	4	US-09-134-001C-4133	Sequence 4133, Ap
452	6	1.3	246	4	US-09-709-791-3	Sequence 3, Appl	525	6	1.3	321	1	US-08-365-103B-8	Sequence 8, Appl
453	6	1.3	249	1	US-08-466-603-2	Sequence 2, Appl	526	6	1.3	324	4	US-09-134-001C-3365	Sequence 3365, Ap
454	6	1.3	249	1	US-08-314-503A-2	Sequence 2, Appl	527	6	1.3	329	4	US-09-134-001C-2889	Sequence 2889, Ap
455	6	1.3	249	1	US-08-468-066-2	Sequence 2, Appl	528	6	1.3	332	4	US-09-134-001C-3199	Sequence 3199, Ap
456	6	1.3	249	2	US-08-466-717-2	Sequence 2, Appl	529	6	1.3	333	1	US-08-148-215A-4	Sequence 4, Appl
457	6	1.3	249	2	US-08-766-738-4	Sequence 4, Appl	530	6	1.3	333	3	US-08-988-876-6	Sequence 6, Appl
458	6	1.3	249	2	US-09-154-802-1	Sequence 1, Appl	531	6	1.3	347	1	US-08-249-420-2	Sequence 2, Appl
459	6	1.3	249	3	US-09-373-029-1	Sequence 1, Appl	532	6	1.3	347	2	US-08-737-663-2	Sequence 2, Appl
460	6	1.3	249	3	US-08-466-743-2	Sequence 2, Appl	533	6	1.3	347	2	US-09-145-828A-16	Sequence 16, Appl
461	6	1.3	249	4	US-09-262-610-4	Sequence 4, Appl	534	6	1.3	348	4	US-08-430-286A-11	Sequence 11, Appl
462	6	1.3	249	5	PCT-US95-12414-2	Sequence 2, Appl	535	6	1.3	348	4	US-09-134-001C-4857	Sequence 4857, Ap
463	6	1.3	251	4	US-08-766-738-3	Sequence 3, Appl	536	6	1.3	350	6	5352575-7	Patent No. 5352575
464	6	1.3	251	4	US-09-262-610-3	Sequence 3, Appl	537	6	1.3	351	3	US-08-688-988-28	Sequence 28, Appl
465	6	1.3	253	2	US-08-474-379C-46	Sequence 46, Appl	538	6	1.3	353	4	US-09-134-001C-4252	Sequence 4252, Ap

539	6	1.3	354	3	US-08-688-988-29	Sequence 29, Appl	612	6	1.3	452	1	US-07-937-609-16	Sequence 16, Appl
540	6	1.3	359	3	US-08-688-988-32	Sequence 32, Appl	613	6	1.3	452	1	US-08-029-170-16	Sequence 16, Appl
541	6	1.3	359	3	US-09-082-089-3	Sequence 3, Appl	614	6	1.3	453	4	US-07-937-609-26	Sequence 26, Appl
542	6	1.3	363	3	US-09-082-089-5	Sequence 5, Appl	615	6	1.3	453	1	US-07-937-609-27	Sequence 27, Appl
543	6	1.3	366	1	US-08-700-359-22	Sequence 22, Appl	616	6	1.3	453	1	US-07-978-892A-5	Sequence 5, Appl
544	6	1.3	369	1	US-08-844-055-2	Sequence 2, Appl	617	6	1.3	453	1	US-08-570-157-4	Sequence 4, Appl
545	6	1.3	369	3	US-09-006-849-2	Sequence 2, Appl	618	6	1.3	453	4	US-08-029-170-26	Sequence 26, Appl
546	6	1.3	370	4	US-09-325-932A-169	Sequence 169, App	619	6	1.3	453	4	US-08-029-170-27	Sequence 27, Appl
547	6	1.3	372	3	US-09-082-089-2	Sequence 2, Appl	620	6	1.3	453	4	US-09-076-510-4	Sequence 4, Appl
548	6	1.3	386	4	US-09-293-322C-5	Sequence 5, Appl	621	6	1.3	457	4	US-08-821-818-1	Sequence 1, Appl
549	6	1.3	386	4	US-09-134-001C-3809	Sequence 3809, Ap	622	6	1.3	457	4	US-08-821-818-5	Sequence 5, Appl
550	6	1.3	390	3	US-08-689-421-21	Sequence 21, Appl	623	6	1.3	463	3	US-09-082-310-1	Sequence 1, Appl
551	6	1.3	390	3	US-08-650-766-7	Sequence 7, Appl	624	6	1.3	463	3	US-09-575-205-1	Sequence 1, Appl
552	6	1.3	390	3	US-08-922-635-6	Sequence 6, Appl	625	6	1.3	465	2	US-08-114-555A-2	Sequence 2, Appl
553	6	1.3	390	4	US-09-389-628-21	Sequence 21, Appl	626	6	1.3	465	2	US-08-833-678A-2	Sequence 2, Appl
554	6	1.3	390	4	US-09-181-827A-21	Sequence 21, Appl	627	6	1.3	465	3	US-08-553-337A-2	Sequence 2, Appl
555	6	1.3	391	4	US-09-134-001C-3442	Sequence 3442, Ap	628	6	1.3	465	4	US-08-529-169A-2	Sequence 2, Appl
556	6	1.3	392	4	US-09-378-088A-90	Sequence 90, Appl	629	6	1.3	470	2	US-08-466-906B-8	Sequence 8, Appl
557	6	1.3	393	1	US-07-629-1041-3	Sequence 3, Appl	630	6	1.3	470	2	US-09-201-746-8	Sequence 8, Appl
558	6	1.3	393	2	US-08-977-554-2	Sequence 2, Appl	631	6	1.3	475	2	US-08-861-464-14	Sequence 14, Appl
559	6	1.3	393	4	US-09-225-967-2	Sequence 2, Appl	632	6	1.3	475	2	US-08-396-001-14	Sequence 14, Appl
560	6	1.3	393	4	US-09-227-806-2	Sequence 2, Appl	633	6	1.3	475	2	US-09-323-433A-14	Sequence 14, Appl
561	6	1.3	397	4	US-09-459-133-2	Sequence 2, Appl	634	6	1.3	479	4	US-08-416-788-2	Sequence 2, Appl
562	6	1.3	398	2	US-08-288-663A-15	Sequence 15, Appl	635	6	1.3	481	1	US-08-286-856C-2	Sequence 2, Appl
563	6	1.3	400	1	US-07-989-991A-2	Sequence 2, Appl	636	6	1.3	481	1	US-08-477-831-2	Sequence 2, Appl
564	6	1.3	400	3	US-08-621-255-2	Sequence 2, Appl	637	6	1.3	486	3	US-08-904-852-2	Sequence 2, Appl
565	6	1.3	400	4	US-09-352-574-2	Sequence 2, Appl	638	6	1.3	486	4	US-08-259-451-13	Sequence 13, Appl
566	6	1.3	402	1	US-08-553-703A-1	Sequence 1, Appl	639	6	1.3	486	4	US-09-517-639-2	Sequence 2, Appl
567	6	1.3	402	1	US-08-553-703A-4	Sequence 4, Appl	640	6	1.3	493	2	US-09-031-392-10	Sequence 10, Appl
568	6	1.3	402	2	US-09-006-021-1	Sequence 1, Appl	641	6	1.3	493	4	US-09-299-549-10	Sequence 10, Appl
569	6	1.3	402	2	US-09-006-021-4	Sequence 4, Appl	642	6	1.3	493	4	US-09-610-417-10	Sequence 10, Appl
570	6	1.3	406	1	US-08-434-881-2	Sequence 2, Appl	643	6	1.3	494	4	US-09-134-001C-1176	Sequence 4176, Ap
571	6	1.3	406	3	US-08-977-771-2	Sequence 2, Appl	644	6	1.3	496	4	US-09-350-268-2	Sequence 2, Appl
572	6	1.3	406	4	US-09-361-773-2	Sequence 2, Appl	645	6	1.3	496	4	US-09-536-679-2	Sequence 2, Appl
573	6	1.3	406	4	US-09-336-536-23	Sequence 23, Appl	646	6	1.3	496	2	US-08-702-598-2	Sequence 2, Appl
574	6	1.3	410	6	5177197-1	Patent No. 5177197	647	6	1.3	498	4	US-09-134-607A-17	Sequence 17, Appl
575	6	1.3	414	1	US-07-667-276A-4	Sequence 4, Appl	648	6	1.3	498	4	US-09-134-607A-18	Sequence 18, Appl
576	6	1.3	415	4	US-08-795-430-11	Sequence 11, Appl	649	6	1.3	498	4	US-09-134-607A-19	Sequence 19, Appl
577	6	1.3	415	4	US-09-355-700-11	Sequence 11, Appl	650	6	1.3	500	1	US-07-755-573C-8	Sequence 8, Appl
578	6	1.3	415	4	US-08-601-132-41	Sequence 41, Appl	651	6	1.3	500	2	US-08-314-601-2	Sequence 2, Appl
579	6	1.3	418	1	US-07-816-283-10	Sequence 10, Appl	652	6	1.3	500	2	PCT-US95-13051-2	Sequence 2, Appl
580	6	1.3	418	1	US-08-417-103-10	Sequence 10, Appl	653	6	1.3	501	2	US-08-577-432A-1	Sequence 40, Appl
581	6	1.3	419	1	US-08-615-170-16	Sequence 16, Appl	654	6	1.3	501	2	US-08-408-095-31	Sequence 31, Appl
582	6	1.3	419	4	US-09-134-001C-3441	Sequence 3441, Ap	655	6	1.3	501	4	US-09-079-630-40	Sequence 40, Appl
583	6	1.3	421	3	US-09-002-567B-1	Sequence 1, Appl	656	6	1.3	505	1	US-09-041-075A-8	Sequence 8, Appl
584	6	1.3	421	3	US-09-002-567B-3	Sequence 3, Appl	657	6	1.3	506	1	US-08-286-856C-3	Sequence 3, Appl
585	6	1.3	421	4	US-09-020-846-68	Sequence 68, Appl	658	6	1.3	506	1	US-08-472-831-3	Sequence 3, Appl
586	6	1.3	421	4	US-09-571-347-1	Sequence 1, Appl	659	6	1.3	508	2	US-08-850-293-5	Sequence 5, Appl
587	6	1.3	421	4	US-09-571-347-3	Sequence 3, Appl	660	6	1.3	513	4	US-08-867-611-59	Sequence 59, Appl
588	6	1.3	424	4	US-09-134-001C-5009	Sequence 5009, Ap	661	6	1.3	517	3	US-08-688-421-33	Sequence 33, Appl
589	6	1.3	426	1	US-08-615-170-6	Sequence 6, Appl	662	6	1.3	517	4	US-09-388-558-33	Sequence 33, Appl
590	6	1.3	426	1	US-08-615-170-15	Sequence 15, Appl	663	6	1.3	517	4	US-09-181-827A-33	Sequence 33, Appl
591	6	1.3	428	1	US-07-816-283-12	Sequence 12, Appl	664	6	1.3	521	4	US-09-258-349-2	Sequence 2, Appl
592	6	1.3	429	1	US-08-906-744A-2	Sequence 2, Appl	665	6	1.3	531	4	US-08-333-152A-2	Sequence 32, Appl
593	6	1.3	429	3	US-09-093-134-2	Sequence 2, Appl	666	6	1.3	534	4	US-09-029-348-5	Sequence 5, Appl
594	6	1.3	431	1	US-08-190-802A-37	Sequence 37, Appl	667	6	1.3	535	4	US-09-029-348-1	Sequence 1, Appl
595	6	1.3	431	4	US-08-477-346-37	Sequence 37, Appl	668	6	1.3	537	4	US-09-029-348-4	Sequence 4, Appl
596	6	1.3	431	4	US-08-473-089-37	Sequence 37, Appl	669	6	1.3	538	4	US-09-647-390-16	Sequence 16, Appl
597	6	1.3	431	4	US-08-487-072A-37	Sequence 37, Appl	670	6	1.3	542	1	US-08-140-729A-5	Sequence 5, Appl
598	6	1.3	437	4	US-09-134-001C-4808	Sequence 4808, Ap	671	6	1.3	542	1	US-08-701-380-2	Sequence 2, Appl
599	6	1.3	437	2	US-08-577-492-38	Sequence 38, Appl	672	6	1.3	542	1	US-08-546-666-5	Sequence 5, Appl
600	6	1.3	438	4	US-09-079-630-38	Sequence 38, Appl	673	6	1.3	542	2	US-08-916-745-5	Sequence 5, Appl
601	6	1.3	438	4	US-09-339-838-5	Sequence 5, Appl	674	6	1.3	542	2	US-08-948-569A-2	Sequence 2, Appl
602	6	1.3	438	4	US-09-339-838-7	Sequence 7, Appl	675	6	1.3	542	2	US-08-663-808-4	Sequence 2, Appl
603	6	1.3	441	4	US-09-522-666-4	Sequence 4, Appl	676	6	1.3	542	2	US-09-042-929-5	Sequence 5, Appl
604	6	1.3	441	1	US-07-937-609-29	Sequence 29, Appl	677	6	1.3	542	2	US-08-546-661-5	Sequence 5, Appl
605	6	1.3	447	1	US-07-978-892A-6	Sequence 6, Appl	678	6	1.3	542	2	US-09-042-929-5	Sequence 5, Appl
606	6	1.3	447	1	US-08-029-170-29	Sequence 29, Appl	679	6	1.3	542	2	US-09-188-469-2	Sequence 2, Appl
607	6	1.3	448	1	US-08-570-157-3	Sequence 3, Appl	680	6	1.3	542	3	US-09-198-650-5	Sequence 5, Appl
608	6	1.3	448	1	US-09-076-510-3	Sequence 3, Appl	681	6	1.3	542	3	US-09-333-740-4	Sequence 4, Appl
609	6	1.3	451	4	US-08-570-157-2	Sequence 2, Appl	682	6	1.3	542	3	US-09-042-913-5	Sequence 5, Appl
610	6	1.3	451	4	US-09-076-510-2	Sequence 2, Appl	683	6	1.3	542	3	US-09-188-496-4	Sequence 4, Appl
611	6	1.3	451	4	US-09-184-964-3	Sequence 3, Appl	684	6	1.3	542	3	US-09-042-937-5	Sequence 5, Appl

685	6	1.3	542	3	US-09-032-365A-13	Sequence 13, Appl	758	6	1.3	609	4	US-09-315-850-40	Sequence 40, Appl
686	6	1.3	542	4	US-09-397-238A-2	Sequence 2, Appl	759	6	1.3	631	1	US-08-700-356-1	Sequence 1, Appl
687	6	1.3	542	4	US-09-368-282-4	Sequence 4, Appl	760	6	1.3	631	2	US-08-936-865-1	Sequence 1, Appl
688	6	1.3	542	4	US-09-566-708A-4	Sequence 4, Appl	761	6	1.3	631	2	US-08-833-678A-1	Sequence 1, Appl
689	6	1.3	542	4	US-09-042-709A-5	Sequence 5, Appl	762	6	1.3	631	2	US-09-128-31A-2	Sequence 2, Appl
690	6	1.3	543	2	US-08-469-412A-7	Sequence 7, Appl	763	6	1.3	631	4	US-08-529-169A-1	Sequence 1, Appl
691	6	1.3	543	2	US-09-021-715-7	Sequence 7, Appl	764	6	1.3	632	4	US-09-198-723A-23	Sequence 23, Appl
692	6	1.3	544	2	US-09-042-709A-18	Sequence 18, Appl	765	6	1.3	646	4	US-09-198-723A-60	Sequence 60, Appl
693	6	1.3	544	2	US-08-935-760-2	Sequence 2, Appl	766	6	1.3	646	4	US-09-198-723A-63	Sequence 63, Appl
694	6	1.3	544	3	US-08-559-397A-19	Sequence 19, Appl	767	6	1.3	646	4	US-09-198-723A-66	Sequence 66, Appl
695	6	1.3	544	3	US-08-459-397A-29	Sequence 29, Appl	768	6	1.3	646	4	US-09-198-723A-69	Sequence 69, Appl
696	6	1.3	548	2	US-08-469-412A-2	Sequence 2, Appl	769	6	1.3	646	4	US-09-198-723A-72	Sequence 72, Appl
697	6	1.3	548	2	US-09-032-315-4	Sequence 4, Appl	770	6	1.3	651	3	US-08-650-766-6	Sequence 6, Appl
698	6	1.3	548	2	US-08-993-318A-4	Sequence 4, Appl	771	6	1.3	651	3	US-08-766-6	Sequence 6, Appl
699	6	1.3	548	4	US-09-399-886-4	Sequence 4, Appl	772	6	1.3	656	2	US-08-533-669A-6	Sequence 6, Appl
700	6	1.3	548	4	US-09-396-260-4	Sequence 4, Appl	773	6	1.3	656	4	US-09-183-861-6	Sequence 6, Appl
701	6	1.3	548	4	US-09-021-715-2	Sequence 2, Appl	774	6	1.3	656	4	US-09-022-765-6	Sequence 6, Appl
702	6	1.3	549	4	US-08-576-281-4	Sequence 4, Appl	775	6	1.3	658	1	US-08-190-802A-34	Sequence 34, Appl
703	6	1.3	553	2	US-08-828-109A-3	Sequence 2, Appl	776	6	1.3	658	4	US-08-477-346-34	Sequence 34, Appl
704	6	1.3	553	2	US-08-943-087-2	Sequence 2, Appl	777	6	1.3	658	4	US-08-473-089-34	Sequence 34, Appl
705	6	1.3	553	2	US-08-943-087-14	Sequence 14, Appl	778	6	1.3	658	4	US-08-487-072A-34	Sequence 34, Appl
706	6	1.3	553	2	US-08-943-087-16	Sequence 16, Appl	779	6	1.3	658	4	US-09-328-599A-2	Sequence 34, Appl
707	6	1.3	553	2	US-08-943-087-18	Sequence 18, Appl	780	6	1.3	661	1	US-08-514-014-4	Sequence 2, Appl
708	6	1.3	553	2	US-08-943-087-20	Sequence 20, Appl	781	6	1.3	661	1	US-08-233-538-12	Sequence 12, Appl
709	6	1.3	553	2	US-08-943-087-22	Sequence 22, Appl	782	6	1.3	661	1	US-08-786-164-12	Sequence 12, Appl
710	6	1.3	553	2	US-08-943-087-24	Sequence 24, Appl	783	6	1.3	661	2	US-08-833-823-4	Sequence 4, Appl
711	6	1.3	553	2	US-08-943-087-26	Sequence 26, Appl	784	6	1.3	666	2	US-08-737-716-14	Sequence 14, Appl
712	6	1.3	553	2	US-08-943-087-28	Sequence 28, Appl	785	6	1.3	666	4	US-09-198-723A-11	Sequence 11, Appl
713	6	1.3	553	2	US-08-943-087-30	Sequence 30, Appl	786	6	1.3	666	4	US-09-198-723A-13	Sequence 13, Appl
714	6	1.3	553	2	US-08-943-087-32	Sequence 32, Appl	787	6	1.3	666	4	US-09-198-723A-14	Sequence 14, Appl
715	6	1.3	553	2	US-08-943-087-34	Sequence 34, Appl	788	6	1.3	666	4	US-09-198-723A-15	Sequence 15, Appl
716	6	1.3	553	2	US-08-943-087-36	Sequence 36, Appl	789	6	1.3	666	4	US-09-198-723A-16	Sequence 16, Appl
717	6	1.3	553	2	US-08-943-087-38	Sequence 38, Appl	790	6	1.3	666	4	US-09-198-723A-17	Sequence 17, Appl
718	6	1.3	553	2	US-08-943-087-40	Sequence 40, Appl	791	6	1.3	666	4	US-09-198-723A-18	Sequence 18, Appl
719	6	1.3	553	2	US-08-943-087-42	Sequence 42, Appl	792	6	1.3	667	2	US-08-718-661-2	Sequence 2, Appl
720	6	1.3	553	2	US-08-943-087-44	Sequence 44, Appl	793	6	1.3	667	2	US-08-718-661-2	Sequence 2, Appl
721	6	1.3	553	2	US-08-943-087-46	Sequence 46, Appl	794	6	1.3	668	4	US-09-134-001C-4816	Sequence 4816, Ap
722	6	1.3	553	2	US-08-943-087-48	Sequence 48, Appl	795	6	1.3	670	2	US-08-366-547-2	Sequence 2, Appl
723	6	1.3	556	1	US-08-810-116-11	Sequence 11, Appl	796	6	1.3	672	4	US-09-292-858B-2	Sequence 2, Appl
724	6	1.3	566	1	US-07-930-548A-11	Sequence 11, Appl	797	6	1.3	672	4	US-09-198-723A-19	Sequence 19, Appl
725	6	1.3	571	1	US-07-955-905A-25	Sequence 25, Appl	798	6	1.3	672	4	US-09-198-723A-20	Sequence 20, Appl
726	6	1.3	573	1	US-09-042-709A-19	Sequence 19, Appl	799	6	1.3	674	3	US-08-893-852A-1	Sequence 1, Appl
727	6	1.3	574	1	US-08-140-729A-7	Sequence 7, Appl	800	6	1.3	675	1	US-08-317-522A-9	Sequence 9, Appl
728	6	1.3	574	1	US-08-546-666-7	Sequence 7, Appl	801	6	1.3	675	1	US-08-439-818A-9	Sequence 9, Appl
729	6	1.3	574	2	US-08-916-745-7	Sequence 7, Appl	802	6	1.3	675	2	US-08-751-965-9	Sequence 9, Appl
730	6	1.3	574	2	US-08-948-569A-4	Sequence 4, Appl	803	6	1.3	675	2	US-08-738-975-9	Sequence 9, Appl
731	6	1.3	574	2	US-08-663-808-6	Sequence 6, Appl	804	6	1.3	675	2	US-08-728-626-9	Sequence 9, Appl
732	6	1.3	574	2	US-09-042-929-7	Sequence 7, Appl	805	6	1.3	675	2	US-08-808-599A-9	Sequence 9, Appl
733	6	1.3	574	2	US-08-546-661-7	Sequence 7, Appl	806	6	1.3	679	1	US-08-214-583-2	Sequence 2, Appl
734	6	1.3	574	2	US-09-042-960-7	Sequence 7, Appl	807	6	1.3	686	1	US-08-350-884A-70	Sequence 70, Appl
735	6	1.3	574	2	US-08-906-713-2	Sequence 2, Appl	808	6	1.3	686	2	US-08-709-173-70	Sequence 70, Appl
736	6	1.3	574	2	US-09-188-469-4	Sequence 4, Appl	809	6	1.3	687	1	US-08-709-177-70	Sequence 70, Appl
737	6	1.3	574	3	US-09-332-740-6	Sequence 6, Appl	810	6	1.3	687	2	US-08-232-538-6	Sequence 6, Appl
738	6	1.3	574	3	US-09-042-913-7	Sequence 7, Appl	811	6	1.3	687	2	US-08-786-164-6	Sequence 6, Appl
739	6	1.3	574	3	US-09-188-469-6	Sequence 6, Appl	812	6	1.3	687	4	US-09-427-353-2	Sequence 2, Appl
740	6	1.3	574	3	US-09-042-913-7	Sequence 7, Appl	813	6	1.3	689	4	US-09-232-200-73	Sequence 73, Appl
741	6	1.3	574	3	US-09-042-937-7	Sequence 7, Appl	814	6	1.3	689	4	US-09-232-201-73	Sequence 73, Appl
742	6	1.3	574	4	US-09-397-238A-4	Sequence 4, Appl	815	6	1.3	689	4	US-09-232-201-73	Sequence 73, Appl
743	6	1.3	574	4	US-09-368-282-6	Sequence 6, Appl	816	6	1.3	705	1	US-08-178-477B-2	Sequence 2, Appl
744	6	1.3	574	4	US-09-566-708A-6	Sequence 6, Appl	817	6	1.3	705	4	US-09-370-838-186	Sequence 186, Appl
745	6	1.3	574	4	US-09-042-709A-7	Sequence 7, Appl	818	6	1.3	706	1	US-08-339-152A-29	Sequence 29, Appl
746	6	1.3	582	4	US-09-134-001C-4730	Sequence 4730, Ap	819	6	1.3	706	2	US-08-007-999B-4	Sequence 4, Appl
747	6	1.3	582	4	US-08-850-328-3	Sequence 3, Appl	820	6	1.3	706	2	US-08-689-276A-4	Sequence 4, Appl
748	6	1.3	582	4	US-09-197-814-3	Sequence 3, Appl	821	6	1.3	730	2	US-08-696-944-2	Sequence 2, Appl
749	6	1.3	604	4	US-09-134-001C-4425	Sequence 4425, Ap	822	6	1.3	739	4	US-08-444-818-148	Sequence 148, App
750	6	1.3	606	3	US-08-577-492-32	Sequence 32, Appl	823	6	1.3	744	3	US-08-179-481-2	Sequence 2, Appl
751	6	1.3	606	3	US-08-620-077B-3	Sequence 32, Appl	824	6	1.3	746	3	US-08-434-000A-4	Sequence 4, Appl
752	6	1.3	606	4	US-09-079-630-32	Sequence 32, Appl	825	6	1.3	746	3	US-09-312-157-4	Sequence 4, Appl
753	6	1.3	608	4	US-09-095-385-4	Sequence 4, Appl	826	6	1.3	758	2	US-08-874-678-1	Sequence 1, Appl
754	6	1.3	608	4	US-09-134-001C-4751	Sequence 4751, Ap	827	6	1.3	758	3	US-08-643-839-1	Sequence 1, Appl
755	6	1.3	609	1	US-08-324-977-40	Sequence 40, Appl	828	6	1.3	758	3	US-09-051-363-24	Sequence 24, Appl
756	6	1.3	609	1	US-08-384-616-40	Sequence 40, Appl	829	6	1.3	758	4	US-09-348-886-1	Sequence 1, Appl
757	6	1.3	609	2	US-08-904-686A-40	Sequence 40, Appl	830	6	1.3	758	4	US-09-348-886-1	Sequence 1, Appl



831	6	1.3	763	1	US-08-155-331-13	Sequence 13, Appl	904	6	1.3	1106	1	US-08-477-329-2	Sequence 2, Appl
832	6	1.3	763	1	US-08-424-022-13	Sequence 13, Appl	905	6	1.3	1106	2	US-08-475-458-2	Sequence 2, Appl
833	6	1.3	763	2	US-08-424-017B-13	Sequence 13, Appl	906	6	1.3	1106	2	US-08-460-510-2	Sequence 2, Appl
834	6	1.3	763	5	PCT-US93-11696-13	Sequence 13, Appl	907	6	1.3	1106	2	US-08-460-510-2	Sequence 2, Appl
835	6	1.3	780	1	US-08-232-358-14	Sequence 14, Appl	908	6	1.3	1106	3	US-08-980-400-2	Sequence 4, Appl
836	6	1.3	780	2	US-08-786-164-14	Sequence 14, Appl	909	6	1.3	1106	3	US-08-462-728-4	Sequence 4, Appl
837	6	1.3	781	4	US-08-867-611-4	Sequence 4, Appl	910	6	1.3	1106	4	US-09-583-459A-2	Sequence 2, Appl
838	6	1.3	781	4	US-09-766-387-2	Sequence 4, Appl	911	6	1.3	1106	4	US-09-583-210-2	Sequence 2, Appl
839	6	1.3	781	4	US-09-766-387-4	Sequence 4, Appl	912	6	1.3	1106	4	US-09-583-449A-2	Sequence 2, Appl
840	6	1.3	781	5	PCT-US92-06865A-9	Sequence 9, Appl	913	6	1.3	1106	4	US-09-433-039-2	Sequence 2, Appl
841	6	1.3	784	4	US-09-740-235-1	Sequence 1, Appl	914	6	1.3	1106	4	US-08-461-917-4	Sequence 4, Appl
842	6	1.3	790	2	US-08-359-705B-9	Sequence 9, Appl	915	6	1.3	1106	5	PCT-US92-00730-2	Sequence 2, Appl
843	6	1.3	790	2	US-08-286-846A-9	Sequence 9, Appl	916	6	1.3	1106	5	PCT-US92-00862-2	Sequence 2, Appl
844	6	1.3	790	2	US-08-457-880A-9	Sequence 9, Appl	917	6	1.3	1135	2	US-08-574-959A-7	Sequence 7, Appl
845	6	1.3	790	3	US-08-444-622A-9	Sequence 9, Appl	918	6	1.3	1135	4	US-09-357-014-7	Sequence 7, Appl
846	6	1.3	790	3	US-08-942-562-9	Sequence 9, Appl	919	6	1.3	1145	4	US-09-470-443-2	Sequence 4, Appl
847	6	1.3	790	4	US-09-156-923-9	Sequence 9, Appl	920	6	1.3	1145	4	US-09-470-443-4	Sequence 4, Appl
848	6	1.3	829	4	US-08-444-818-69	Sequence 69, Appl	921	6	1.3	1178	1	US-08-199-776-2	Sequence 2, Appl
849	6	1.3	829	4	US-09-562-737-40	Sequence 40, Appl	922	6	1.3	1178	3	US-08-663-731-2	Sequence 2, Appl
850	6	1.3	830	4	US-09-562-737-32	Sequence 32, Appl	923	6	1.3	1178	3	US-08-879-338-2	Sequence 2, Appl
851	6	1.3	830	4	US-09-562-737-37	Sequence 37, Appl	924	6	1.3	1178	5	PCT-US95-02044-2	Sequence 2, Appl
852	6	1.3	830	4	US-09-562-737-39	Sequence 39, Appl	925	6	1.3	1179	4	US-09-293-228B-2	Sequence 2, Appl
853	6	1.3	841	1	US-08-350-884-86	Sequence 86, Appl	926	6	1.3	1194	4	US-08-538-526-1	Sequence 4, Appl
854	6	1.3	841	1	US-08-709-173-86	Sequence 86, Appl	927	6	1.3	1196	4	US-08-881-706-2	Sequence 2, Appl
855	6	1.3	841	2	US-08-709-177-86	Sequence 86, Appl	928	6	1.3	1229	3	US-09-310-293-2	Sequence 2, Appl
856	6	1.3	859	2	US-08-444-818-30	Sequence 30, Appl	929	6	1.3	1244	5	US-09-579-376-2	Sequence 2, Appl
857	6	1.3	865	1	US-07-803-633A-13	Sequence 13, Appl	930	6	1.3	1244	5	PCT-US93-10500-2	Sequence 2, Appl
858	6	1.3	866	4	US-09-556-877-189	Sequence 189, App	931	6	1.3	1248	4	US-08-882-046-6	Sequence 6, Appl
859	6	1.3	866	4	US-09-620-412C-189	Sequence 189, App	932	6	1.3	1257	4	US-09-220-641-3	Sequence 3, Appl
860	6	1.3	873	4	US-09-540-824-28	Sequence 28, Appl	933	6	1.3	1259	4	US-09-134-001C-3757	Sequence 3757, Ap
861	6	1.3	880	4	US-09-556-877-175	Sequence 175, App	934	6	1.3	1311	1	US-08-340-011-5	Sequence 5, Appl
862	6	1.3	880	4	US-09-620-412C-175	Sequence 175, App	935	6	1.3	1311	3	US-08-901-710-5	Sequence 5, Appl
863	6	1.3	905	1	US-08-072-574-2	Sequence 2, Appl	936	6	1.3	1332	2	US-08-927-244-2	Sequence 2, Appl
864	6	1.3	905	2	US-08-574-959A-9	Sequence 9, Appl	937	6	1.3	1332	4	US-09-286-831-2	Sequence 2, Appl
865	6	1.3	905	4	US-09-357-014-9	Sequence 9, Appl	938	6	1.3	1338	4	US-08-750-141A-3	Sequence 3, Appl
866	6	1.3	906	1	US-08-486-270-2	Sequence 2, Appl	939	6	1.3	1362	2	US-08-874-678-33	Sequence 33, Appl
867	6	1.3	906	3	US-08-367-264-2	Sequence 2, Appl	940	6	1.3	1362	3	US-08-643-839-33	Sequence 33, Appl
868	6	1.3	906	4	US-09-153-757-2	Sequence 2, Appl	941	6	1.3	1362	4	US-09-348-886-33	Sequence 33, Appl
869	6	1.3	907	3	US-08-783-774-2	Sequence 2, Appl	942	6	1.3	1366	3	US-08-963-825-19	Sequence 19, Appl
870	6	1.3	907	4	US-09-328-559A-1	Sequence 1, Appl	943	6	1.3	1366	4	US-09-500-811-19	Sequence 19, Appl
871	6	1.3	907	5	PCT-US95-04611A-19	Sequence 19, Appl	944	6	1.3	1366	4	US-09-570-678-19	Sequence 19, Appl
872	6	1.3	913	1	US-08-445-640-4	Sequence 4, Appl	945	6	1.3	1366	4	US-09-548-608-19	Sequence 19, Appl
873	6	1.3	913	3	US-08-170-558-4	Sequence 4, Appl	946	6	1.3	1366	4	US-09-585-887-10	Sequence 10, Appl
874	6	1.3	913	3	US-08-447-314-4	Sequence 4, Appl	947	6	1.3	1366	4	US-09-289-578-10	Sequence 10, Appl
875	6	1.3	913	3	US-08-445-461-4	Sequence 4, Appl	948	6	1.3	1394	6	5177197-30	Patent No. 5177197
876	6	1.3	914	1	US-08-484-105-2	Sequence 2, Appl	949	6	1.3	1420	2	US-08-540-804-14	Sequence 14, Appl
877	6	1.3	914	1	US-08-484-106-2	Sequence 2, Appl	950	6	1.3	1420	2	US-08-218-265-14	Sequence 14, Appl
878	6	1.3	918	2	US-08-843-530B-35	Sequence 35, Appl	951	6	1.3	1420	3	US-08-521-872-14	Sequence 14, Appl
879	6	1.3	919	1	US-08-336-343A-2	Sequence 2, Appl	952	6	1.3	1420	4	US-08-559-399-14	Sequence 14, Appl
880	6	1.3	967	1	US-08-188-281B-13	Sequence 13, Appl	953	6	1.3	1648	1	US-08-188-281B-12	Sequence 12, Appl
881	6	1.3	967	4	US-09-130-481-2	Sequence 2, Appl	954	6	1.3	1648	5	PCT-US94-07280-12	Sequence 12, Appl
882	6	1.3	967	5	PCT-US94-07280-13	Sequence 13, Appl	955	6	1.3	1648	5	PCT-US95-01087-12	Sequence 12, Appl
883	6	1.3	967	5	PCT-US95-01087-13	Sequence 13, Appl	956	6	1.3	1676	4	US-08-487-283A-2	Sequence 2, Appl
884	6	1.3	971	4	US-08-867-611-52	Sequence 52, Appl	957	6	1.3	1692	4	US-09-263-933-4	Sequence 4, Appl
885	6	1.3	973	3	US-08-904-452-4	Sequence 4, Appl	958	6	1.3	1692	4	US-09-263-933-11	Sequence 11, Appl
886	6	1.3	973	3	US-08-867-611-53	Sequence 53, Appl	959	6	1.3	1692	4	US-09-263-933-18	Sequence 18, Appl
887	6	1.3	973	4	US-09-517-639-4	Sequence 4, Appl	960	6	1.3	1719	2	US-08-455-568-4	Sequence 4, Appl
888	6	1.3	992	4	US-08-867-611-54	Sequence 54, Appl	961	6	1.3	1719	2	US-08-399-411-4	Sequence 4, Appl
889	6	1.3	1021	1	US-08-497-025-3	Sequence 3, Appl	962	6	1.3	1719	3	US-08-516-859A-4	Sequence 4, Appl
890	6	1.3	1021	1	US-07-910-760-12	Sequence 12, Appl	963	6	1.3	1719	4	US-09-586-472-4	Sequence 4, Appl
891	6	1.3	1021	1	US-08-440-519-12	Sequence 12, Appl	964	6	1.3	1719	4	US-09-528-706-4	Sequence 4, Appl
892	6	1.3	1021	4	US-08-440-549-12	Sequence 12, Appl	965	6	1.3	1732	2	US-08-477-451-14	Sequence 14, Appl
893	6	1.3	1036	2	US-08-720-484A-5	Sequence 5, Appl	966	6	1.3	1786	4	US-08-444-818-54	Sequence 54, Appl
894	6	1.3	1036	4	US-08-953-833A-5	Sequence 5, Appl	967	6	1.3	1872	4	US-08-188-582-14	Sequence 14, Appl
895	6	1.3	1036	4	US-09-398-239-5	Sequence 5, Appl	968	6	1.3	1872	1	US-08-646-715-14	Sequence 14, Appl
896	6	1.3	1044	2	US-08-777-405A-2	Sequence 2, Appl	969	6	1.3	1893	1	US-08-188-582-11	Sequence 11, Appl
897	6	1.3	1044	2	US-08-977-871A-2	Sequence 2, Appl	970	6	1.3	1893	1	US-08-646-715-11	Sequence 11, Appl
898	6	1.3	1049	3	US-09-225-951-2	Sequence 2, Appl	971	6	1.3	2004	1	US-08-375-709-15	Sequence 15, Appl
899	6	1.3	1049	3	US-08-772-270A-11	Sequence 11, Appl	972	6	1.3	2004	1	US-08-752-929-15	Sequence 15, Appl
900	6	1.3	1070	3	US-08-922-635-22	Sequence 22, Appl	973	6	1.3	2004	4	US-09-090-793-9	Sequence 9, Appl
901	6	1.3	1076	4	US-09-470-443-6	Sequence 6, Appl	974	6	1.3	2013	1	US-08-324-977-12	Sequence 12, Appl
902	6	1.3	1106	1	US-08-180-195-2	Sequence 2, Appl	975	6	1.3	2013	2	US-08-384-616-12	Sequence 12, Appl
903	6	1.3	1106	1	US-08-168-917-2	Sequence 2, Appl	976	6	1.3	2013	2	US-08-904-686A-12	Sequence 12, Appl

977 6 1.3 2013 4 US-09-315-850-12 Sequence 12, Appl  
978 6 1.3 2201 4 US-08-952-981A-2 Sequence 2, Appl  
979 6 1.3 2233 2 US-08-569-853-1 Sequence 1, Appl  
980 6 1.3 2233 2 US-08-569-853-2 Sequence 2, Appl  
981 6 1.3 2233 3 US-08-987-439-1 Sequence 1, Appl  
982 6 1.3 2261 4 US-08-444-818-66 Sequence 66, Appl  
983 6 1.3 2307 4 US-09-263-933-2 Sequence 2, Appl  
984 6 1.3 2307 4 US-09-263-933-2 Sequence 9, Appl  
985 6 1.3 2307 4 US-09-263-933-16 Sequence 16, Appl  
986 6 1.3 2318 4 US-09-091-219-24 Sequence 24, Appl  
987 6 1.3 2353 4 US-08-984-709A-50 Sequence 50, Appl  
988 6 1.3 2436 4 US-08-444-818-75 Sequence 75, Appl  
989 6 1.3 2476 2 US-08-324-977-32 Sequence 2, Appl  
990 6 1.3 2620 2 US-08-384-616-32 Sequence 32, Appl  
991 6 1.3 2620 2 US-08-904-686A-32 Sequence 32, Appl  
992 6 1.3 2620 2 US-09-315-850-32 Sequence 32, Appl  
993 6 1.3 2621 1 US-08-324-977-36 Sequence 36, Appl  
994 6 1.3 2621 2 US-08-384-616-36 Sequence 36, Appl  
995 6 1.3 2621 2 US-09-315-850-36 Sequence 36, Appl  
996 6 1.3 2621 4 US-08-375-709-11 Sequence 11, Appl  
997 6 1.3 2756 1 US-08-752-929-11 Sequence 11, Appl  
998 6 1.3 2756 1 US-09-090-793-7 Sequence 7, Appl  
1000 6 1.3 2756 4

## ALIGNMENTS

RESULT 1  
US-08-967-101-2

Sequence 2, Application US/08967101

Patent No. 5840540

GENERAL INFORMATION:

APPLICANT: ST. GEORGE-HYSLOP, PETER H

APPLICANT: ROMMENS, JOHANNA M

APPLICANT: FRASER, PAUL E

TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED

TITLE OF INVENTION: TO ALZHEIMER'S DISEASE

NUMBER OF SEQUENCES: 183

CORRESPONDENCE ADDRESS:

ADDRESSEE: TESTA, HURWITZ &amp; THIBEAULT

STREET: High Street Tower - 125 High Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/967,101

FILING DATE: 10-NOV-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/592,541

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Pitcher, Edmund R.

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-967-101-2

Query Match 100.0%; Score 467; DB 2; Length 467;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEELPAPISYFONAOSEEDNHLSTNVRSONDRREOENHRSLSGHPPELSNGROGNSR 60  
1 MEELPAPISYFONAOSEEDNHLSTNVRSONDRREOENHRSLSGHPPELSNGROGNSR 60  
Db 1 MEELPAPISYFONAOSEEDNHLSTNVRSONDRREOENHRSLSGHPPELSNGROGNSR 60

QY 61 QVEDDEDEDELTLYKYGAKHVMFVPTLCMVVVVATIKSVSYTRKDGOLITPPE 120  
61 QVEDDEDEDELTLYKYGAKHVMFVPTLCMVVVVATIKSVSYTRKDGOLITPPE 120  
Db 61 QVEDDEDEDELTLYKYGAKHVMFVPTLCMVVVVATIKSVSYTRKDGOLITPPE 120

QY 121 DTEYGOALHSILNAAIMISYIVMTLLVLYRYCYVIAHMLIISLLFFFSFI 180  
121 DTEYGOALHSILNAAIMISYIVMTLLVLYRYCYVIAHMLIISLLFFFSFI 180  
Db 121 DTEYGOALHSILNAAIMISYIVMTLLVLYRYCYVIAHMLIISLLFFFSFI 180

QY 181 YLGEVFKYNNVADVITYALLIWNLVGVGMISIHKKGPLRLOQATLIMISALMLVFITY 240  
181 YLGEVFKYNNVADVITYALLIWNLVGVGMISIHKKGPLRLOQATLIMISALMLVFITY 240  
Db 181 YLGEVFKYNNVADVITYALLIWNLVGVGMISIHKKGPLRLOQATLIMISALMLVFITY 240

QY 241 LPEWTAMTILAVISYDVAVLCPEKPLRMLVETAOENETLFPALITYSTPMWLNNMAE 300  
241 LPEWTAMTILAVISYDVAVLCPEKPLRMLVETAOENETLFPALITYSTPMWLNNMAE 300  
Db 241 LPEWTAMTILAVISYDVAVLCPEKPLRMLVETAOENETLFPALITYSTPMWLNNMAE 300

QY 301 GDPAQRVRSKNSKYNAESTERESODIVAENDDGFSEWEAORDSHLGPRTSPESRAA 360  
301 GDPAQRVRSKNSKYNAESTERESODIVAENDDGFSEWEAORDSHLGPRTSPESRAA 360  
Db 301 GDPAQRVRSKNSKYNAESTERESODIVAENDDGFSEWEAORDSHLGPRTSPESRAA 360

QY 361 VOELSSSLIAGEDPERKVKLGDFITFYSLVGKASTAGDMNTTACVAILIGLCL 420  
361 VOELSSSLIAGEDPERKVKLGDFITFYSLVGKASTAGDMNTTACVAILIGLCL 420  
Db 361 VOELSSSLIAGEDPERKVKLGDFITFYSLVGKASTAGDMNTTACVAILIGLCL 420

QY 421 TLLLAIFKKALPALPISITFGLVFYFATDLYVOPFMQOLFHOXYI 467  
421 TLLLAIFKKALPALPISITFGLVFYFATDLYVOPFMQOLFHOXYI 467  
Db 421 TLLLAIFKKALPALPISITFGLVFYFATDLYVOPFMQOLFHOXYI 467

RESULT 2  
US-08-592-541-2

Sequence 2, Application US/08592541

Patent No. 5986054

GENERAL INFORMATION:

APPLICANT: ST. GEORGE-HYSLOP, PETER H

APPLICANT: ROMMENS, JOHANNA M

APPLICANT: FRASER, PAUL E

TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED

TITLE OF INVENTION: TO ALZHEIMER'S DISEASE

NUMBER OF SEQUENCES: 183

CORRESPONDENCE ADDRESS:

ADDRESSEE: TESTA, HURWITZ &amp; THIBEAULT

STREET: High Street Tower - 125 High Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/592,541

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Pitcher, Edmund R.

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-592-541-2

## Query Match

Best Local Similarity 100.0%; Score 467; DB 2; Length 467;  
Pred. No. 0;  
Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTELPAPLSTYFQNMQMSDNHLSNTVRSQNDNREQEHNDRLSGHPPLSNGRPGNSR 60  
DB 1 MTELPAPLSTYFQNMQMSDNHLSNTVRSQNDNREQEHNDRLSGHPPLSNGRPGNSR 60  
QY 61 QVQDEDEDEDELTKYGAHVIMLFVPTLCMVVAVATISVSFTYTRKDCOLIYTPETE 120  
DB 61 QVQDEDEDEDELTKYGAHVIMLFVPTLCMVVAVATISVSFTYTRKDCOLIYTPETE 120  
QY 121 DTEYVGRALHSILNAIMISVIYVMTLLVLYKRYCKYIHAWLLISSLLLEFFSFI 180  
DB 121 DTEYVGRALHSILNAIMISVIYVMTLLVLYKRYCKYIHAWLLISSLLLEFFSFI 180  
QY 181 YLGEVFKTYNAVDITYVALLIMNLGVGMISIHKGPLRLQOAYLIMISALMALVFTKY 240  
DB 181 YLGEVFKTYNAVDITYVALLIMNLGVGMISIHKGPLRLQOAYLIMISALMALVFTKY 240  
QY 241 LPETAMILILAVISYDVAVLCPRGRLMLVETAQERNETLFPALISSTWMLVNNAE 300  
DB 241 LPETAMILILAVISYDVAVLCPRGRLMLVETAQERNETLFPALISSTWMLVNNAE 300  
QY 301 GDEPQORRVSKNSKYNAESTERESODTVAENDDGFSEWEQORSHLGPHPSPESRAA 360  
DB 301 GDEPQORRVSKNSKYNAESTERESODTVAENDDGFSEWEQORSHLGPHPSPESRAA 360  
QY 361 VOELSSSILAGEDEPERGVKLGDFIFYSVLVGKASATASGDMWTTTACFAVAILIGLCL 420  
DB 361 VOELSSSILAGEDEPERGVKLGDFIFYSVLVGKASATASGDMWTTTACFAVAILIGLCL 420  
QY 421 TLLLIATFKKALPALPISITFGLVFYFATDYLVPFMDQLAFHOFYI 467  
DB 421 TLLLIATFKKALPALPISITFGLVFYFATDYLVPFMDQLAFHOFYI 467

## RESULT 3

US-09-124-698-2  
Sequence 2, Application US/09124698

## GENERAL INFORMATION:

APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: ROMMENS, JOHANNA M  
APPLICANT: FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
STREET: High Street Tower - 125 High Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/124,698  
FILING DATE:  
CLASSIFICATION:

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/592,541  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:

NAME: Pitcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100

## INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-124-698-2

## Query Match

Best Local Similarity 100.0%; Score 467; DB 3; Length 467;  
Pred. No. 0;  
Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTELPAPLSTYFQNMQMSDNHLSNTVRSQNDNREQEHNDRLSGHPPLSNGRPGNSR 60  
DB 1 MTELPAPLSTYFQNMQMSDNHLSNTVRSQNDNREQEHNDRLSGHPPLSNGRPGNSR 60  
QY 61 QVQDEDEDEDELTKYGAHVIMLFVPTLCMVVAVATISVSFTYTRKDCOLIYTPETE 120  
DB 61 QVQDEDEDEDELTKYGAHVIMLFVPTLCMVVAVATISVSFTYTRKDCOLIYTPETE 120  
QY 121 DTEYVGRALHSILNAIMISVIYVMTLLVLYKRYCKYIHAWLLISSLLLEFFSFI 180  
DB 121 DTEYVGRALHSILNAIMISVIYVMTLLVLYKRYCKYIHAWLLISSLLLEFFSFI 180  
QY 181 YLGEVFKTYNAVDITYVALLIMNLGVGMISIHKGPLRLQOAYLIMISALMALVFTKY 240  
DB 181 YLGEVFKTYNAVDITYVALLIMNLGVGMISIHKGPLRLQOAYLIMISALMALVFTKY 240  
QY 241 LPETAMILILAVISYDVAVLCPRGRLMLVETAQERNETLFPALISSTWMLVNNAE 300  
DB 241 LPETAMILILAVISYDVAVLCPRGRLMLVETAQERNETLFPALISSTWMLVNNAE 300  
QY 301 GDEPQORRVSKNSKYNAESTERESODTVAENDDGFSEWEQORSHLGPHPSPESRAA 360  
DB 301 GDEPQORRVSKNSKYNAESTERESODTVAENDDGFSEWEQORSHLGPHPSPESRAA 360  
QY 361 VOELSSSILAGEDEPERGVKLGDFIFYSVLVGKASATASGDMWTTTACFAVAILIGLCL 420  
DB 361 VOELSSSILAGEDEPERGVKLGDFIFYSVLVGKASATASGDMWTTTACFAVAILIGLCL 420  
QY 421 TLLLIATFKKALPALPISITFGLVFYFATDYLVPFMDQLAFHOFYI 467  
DB 421 TLLLIATFKKALPALPISITFGLVFYFATDYLVPFMDQLAFHOFYI 467

## RESULT 4

US-09-127-480-2  
Sequence 2, Application US/09127480

## GENERAL INFORMATION:

APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: ROMMENS, JOHANNA M  
APPLICANT: FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
STREET: High Street Tower - 125 High Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/127,480  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,541  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pletcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-127-480-2

Query Match 100.0%; Score 467; DB 4; Length 467;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTELPAPLSYFQNAQMSQEDNHLSTVRSQNDNRROEHNDRSLGHEPEPLSNGRPOGNSR 60  
DB 1 MTELPAPLSYFQNAQMSQEDNHLSTVRSQNDNRROEHNDRSLGHEPEPLSNGRPOGNSR 60  
QY 61 QVVEODEDEDELTKYGAHVIMLFVPTLCMVVAVATIKVSFYTARKDQLIYTPETE 120  
DB 61 QVVEODEDEDELTKYGAHVIMLFVPTLCMVVAVATIKVSFYTARKDQLIYTPETE 120  
QY 121 DTEVGGRAHLSILNAIMISVIYVMTLLVLYKRCYKVIHAWLITSSLLFFSF 180  
DB 121 DTEVGGRAHLSILNAIMISVIYVMTLLVLYKRCYKVIHAWLITSSLLFFSF 180  
QY 181 YLGEVFTYNAVVDYITVALLIMNLGVGMISIMKGPRLQQAYLIMISALMALVFITY 240  
DB 181 YLGEVFTYNAVVDYITVALLIMNLGVGMISIMKGPRLQQAYLIMISALMALVFITY 240  
QY 241 LPEWTAMLILAVISYVDLVAVLCPKGPLRMLVETAQERNETLFPALITYSSTMVWLVNMAE 300  
DB 241 LPEWTAMLILAVISYVDLVAVLCPKGPLRMLVETAQERNETLFPALITYSSTMVWLVNMAE 300  
QY 301 GDPFAQRVSKNSKYNAESTERESQDTVAENDDGGFSEEMEAQDSDLGHRSTPESRAA 360  
DB 301 GDPFAQRVSKNSKYNAESTERESQDTVAENDDGGFSEEMEAQDSDLGHRSTPESRAA 360  
QY 361 VOELSSSILAGEDEBERGVKLGDFIFYSVLVKASATASGDMNTTIACFVALITIGLCL 420  
DB 361 VOELSSSILAGEDEBERGVKLGDFIFYSVLVKASATASGDMNTTIACFVALITIGLCL 420  
QY 421 TLLILAIFFKALPALPISITFGVLFYFATDYLVOPFMDQLAFHQFYI 467  
DB 421 TLLILAIFFKALPALPISITFGVLFYFATDYLVOPFMDQLAFHQFYI 467

RESULT 5  
US-08-496-841C-2

Sequence 2, Application US/08496841C  
Patent No. 6210919

GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H

ROMMENS, JOHANNA M

FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TO ALZHEIMER'S DISEASE

NUMBER OF SEQUENCES: 175

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby & Darby, PC  
STREET: 805 Third Avenue  
CITY: New York

STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/496,841C  
FILING DATE: 28-Jun-1995  
CLASSIFICATION: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Paul F. Fehlner, Ph.D.  
REGISTRATION NUMBER: 35,135  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 753-6237  
TELEFAX: (212) 753-6237  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-08-496-841C-2

Query Match 100.0%; Score 467; DB 4; Length 467;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTELPAPLSYFQNAQMSQEDNHLSTVRSQNDNRROEHNDRSLGHEPEPLSNGRPOGNSR 60  
DB 1 MTELPAPLSYFQNAQMSQEDNHLSTVRSQNDNRROEHNDRSLGHEPEPLSNGRPOGNSR 60  
QY 61 QVVEODEDEDELTKYGAHVIMLFVPTLCMVVAVATIKVSFYTARKDQLIYTPETE 120  
DB 61 QVVEODEDEDELTKYGAHVIMLFVPTLCMVVAVATIKVSFYTARKDQLIYTPETE 120  
QY 121 DTEVGGRAHLSILNAIMISVIYVMTLLVLYKRCYKVIHAWLITSSLLFFSF 180  
DB 121 DTEVGGRAHLSILNAIMISVIYVMTLLVLYKRCYKVIHAWLITSSLLFFSF 180  
QY 181 YLGEVFTYNAVVDYITVALLIMNLGVGMISIMKGPRLQQAYLIMISALMALVFITY 240  
DB 181 YLGEVFTYNAVVDYITVALLIMNLGVGMISIMKGPRLQQAYLIMISALMALVFITY 240  
QY 241 LPEWTAMLILAVISYVDLVAVLCPKGPLRMLVETAQERNETLFPALITYSSTMVWLVNMAE 300  
DB 241 LPEWTAMLILAVISYVDLVAVLCPKGPLRMLVETAQERNETLFPALITYSSTMVWLVNMAE 300  
QY 301 GDPFAQRVSKNSKYNAESTERESQDTVAENDDGGFSEEMEAQDSDLGHRSTPESRAA 360  
DB 301 GDPFAQRVSKNSKYNAESTERESQDTVAENDDGGFSEEMEAQDSDLGHRSTPESRAA 360  
QY 361 VOELSSSILAGEDEBERGVKLGDFIFYSVLVKASATASGDMNTTIACFVALITIGLCL 420  
DB 361 VOELSSSILAGEDEBERGVKLGDFIFYSVLVKASATASGDMNTTIACFVALITIGLCL 420  
QY 421 TLLILAIFFKALPALPISITFGVLFYFATDYLVOPFMDQLAFHQFYI 467  
DB 421 TLLILAIFFKALPALPISITFGVLFYFATDYLVOPFMDQLAFHQFYI 467

RESULT 6  
US-09-124-523-2

Sequence 2, Application US/09124523  
Patent No. 6395960

GENERAL INFORMATION:

APPLICANT: ST. GEORGE-HYSLOP, PETER H

ROMMENS, JOHANNA M

FRASER, PAUL E

TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TO ALZHEIMER'S DISEASE  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HUMWITZ & THIBEAULT  
STREET: High Street Tower - 125 High Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent release #1.0, version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/124,523  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,541  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-124-523-2

Query Match 100.0%; Score 467; DB 4; Length 467;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTELPAPLSYFQNAQMSQDNHLSNTVRSQNDNRQEHNDRLSLGHPPLSNGRPGNSR 60  
DB 1 MTELPAPLSYFQNAQMSQDNHLSNTVRSQNDNRQEHNDRLSLGHPPLSNGRPGNSR 60  
QY 61 QVVEDDEDEDELTKYGAHHVIMLFVPTLCMVVVVATIKSVFTRDQGLITPPE 120  
DB 61 QVVEDDEDEDELTKYGAHHVIMLFVPTLCMVVVVATIKSVFTRDQGLITPPE 120  
QY 121 DTEYGORALHSILNAAMISIVVMTLLVLYKRCYKVIHAMLIISSLLFFFSFI 180  
DB 121 DTEYGORALHSILNAAMISIVVMTLLVLYKRCYKVIHAMLIISSLLFFFSFI 180  
QY 181 YLGEVFKTYNVAVDYITVALLIMNLGVGMISIHMKGPLRLOQAVLIMISALMAVFIKY 240  
DB 181 YLGEVFKTYNVAVDYITVALLIMNLGVGMISIHMKGPLRLOQAVLIMISALMAVFIKY 240  
QY 241 LPEWTAAMLILAVISYDLVAVLCPRGPLMLVETAQERNETLFPALIIYSTWMLVNAE 300  
DB 241 LPEWTAAMLILAVISYDLVAVLCPRGPLMLVETAQERNETLFPALIIYSTWMLVNAE 300  
QY 301 GDEPAORRYSKSKYNAESTERESODTVAENDGSGFSEMEQORDSLPHRSTPESRAA 360  
DB 301 GDEPAORRYSKSKYNAESTERESODTVAENDGSGFSEMEQORDSLPHRSTPESRAA 360  
QY 361 VOELSSSILAGEDPEERGVKLGDFIFYSVLGKASATASGDMNTTACFAAILGLCL 420  
DB 361 VOELSSSILAGEDPEERGVKLGDFIFYSVLGKASATASGDMNTTACFAAILGLCL 420  
QY 421 TLLLLAIFKKAIPALPISITFGIVTFATDYLQPFMDLAFHOEYI 467  
DB 421 TLLLLAIFKKAIPALPISITFGIVTFATDYLQPFMDLAFHOEYI 467

RESULT 7  
US-08-888-077A-4  
Sequence 4, Application US/08888077A  
Patent No. 6020143  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: ROMMENS, JOHANNA M  
APPLICANT: FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TO ALZHEIMER'S DISEASE AND USES THEREFOR.  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLTZ & MENTLIK  
STREET: 600 SOUTH AVENUE WEST  
CITY: WESTFIELD  
STATE: NJ  
COUNTRY: USA  
ZIP: 07090-1497  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/888,077A  
FILING DATE: 03-JUL-1997  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/592,541  
FILING DATE: 26-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: PALISI, THOMAS M  
REGISTRATION NUMBER: 36,629  
REFERENCE/DOCKET NUMBER: SCHERING 3.0-017 CIP CIP IV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 654-5000  
TELEFAX: (908) 654-7866  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 463 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-888-077A-4

Query Match 93.8%; Score 438; DB 3; Length 463;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 438; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 NDNREQEHNDRLSLGHPPLSNGRPGNSRQVVEDDEDEDELTKYGAHHVIMLFV 89  
DB 26 NDNREQEHNDRLSLGHPPLSNGRPGNSRQVVEDDEDEDELTKYGAHHVIMLFV 85  
QY 90 TLCMVVVVATIKSVFTRDQGLITPPTEDTETVGORALHSILNAAMISIVVMTLL 149  
DB 86 TLCMVVVVATIKSVFTRDQGLITPPTEDTETVGORALHSILNAAMISIVVMTLL 145  
QY 150 LVLYKRCYKVIHAMLIISSLLFFFSFIYLGVEFKTYNVAVDYITVALLIMNLGV 209  
DB 146 LVLYKRCYKVIHAMLIISSLLFFFSFIYLGVEFKTYNVAVDYITVALLIMNLGV 205  
QY 210 MISIHMKGPLRLOQAVLIMISALMAVFIKYLPREWTAAMLILAVISYDLVAVLCPR 269  
DB 206 MISIHMKGPLRLOQAVLIMISALMAVFIKYLPREWTAAMLILAVISYDLVAVLCPR 265  
QY 270 MLVETAQERNETLFPALIIYSTWMLVNAEGDPEAORRYSKSKYNAESTERESODTVA 329  
DB 266 MLVETAQERNETLFPALIIYSTWMLVNAEGDPEAORRYSKSKYNAESTERESODTVA 325  
QY 330 ENDDGSGFSEMEQORDSLGPHRSTPESRAAVOELSSSILAGEDPEERGVKLGDFIF 389  
DB 326 ENDDGSGFSEMEQORDSLGPHRSTPESRAAVOELSSSILAGEDPEERGVKLGDFIF 385



QY 181 YLGEVFKTYNAVDYITVALLIMNLGVGMISIMHKGPLRLQOAYLIMISALMALVFRTKY 240  
Db 181 YLGEVFKTYNAVDYITVALLIMNLGVGMISIMHKGPLRLQOAYLIMISALMALVFRTKY 240  
QY 241 LPEWTAMLLIAVISYVDLVAVLCPKGPLRMLVETAQERNETLFPALITYSTMVNLVNNAE 300  
Db 241 LPEWTAMLLIAVISYVDLVAVLCPKGPLRMLVETAQERNETLFPALITYSTMVNLVNNAE 300  
QY 301 GDPAQRVRSKNSKYNAESTRESODTYVAENDDGFSEEMEAQRDSDLGPHRSTPESRAA 360  
Db 301 GDPAQRVRSKNSKYNAESTRESODTYVAENDDGFSEEMEAQRDSDLGPHRSTPESRAA 360  
QY 361 VOELSSSILAGEDPEERGVKLGDFIFYSVLGKASATASGDMNTTACFVAILIGLCL 420  
Db 361 VOELSSSILAGEDPEERGVKLGDFIFYSVLGKASATASGDMNTTACFVAILIGLCL 420  
QY 421 TLLLLAIFKKALPALPISITFGLVFYFATDYLVOFMDQLAFHOFYI 467  
Db 421 TLLLLAIFKKALPALPISITFGLVFYFATDYLVOFMDQLAFHOFYI 467

RESULT 10  
US-08-923-454A-10  
: Sequence 10, Application US/08923454A  
: Patent No. 6004794  
: GENERAL INFORMATION:  
: APPLICANT: Creasy, Caretha  
: APPLICANT: Liyi, George  
: APPLICANT: Kairan, Eric  
: APPLICANT: Clinkenbeard, Helen  
: APPLICANT: Browne, Michael  
: APPLICANT: Southan, Christopher  
: TITLE OF INVENTION: HUMAN SERINE PROTEASE  
: NUMBER OF SEQUENCES: 40  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: SmithKline Beecham Corporation  
: STREET: 709 Swedeland Road  
: CITY: King of Prussia  
: STATE: PA  
: COUNTRY: USA  
: ZIP: 19406  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Diskette  
: OPERATING SYSTEM: IBM Compatible  
: SOFTWARE: FASTSEQ Version 1.5  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/923,454A  
: FILING DATE:  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: 60/025436  
: FILING DATE: 06-SEPT-1996  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Baumelster, Kirk  
: REGISTRATION NUMBER: 33,833  
: REFERENCE/DOCKET NUMBER: P50547  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 610-270-5096  
: TELEFAX: 610-270-5090  
: TELEX:  
: INFORMATION FOR SRO ID NO: 10:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 467 amino acids  
: TYPE: amino acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: peptide  
: HYPOTHEICAL: NO  
: ANTI-SENSE: NO  
: FRAGMENT TYPE: N-terminal  
: ORIGINAL SOURCE:

US-08-923-454A-10

Query Match 78.4%; Score 366; DB 3; Length 467;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTELPAPLSTYFQNMOMSDNLSNTVRSQNDNREOEHNDRSRIGHPPLSNGRPOGNSR 60  
Db 1 MTELPAPLSTYFQNMOMSDNLSNTVRSQNDNREOEHNDRSRIGHPPLSNGRPOGNSR 60  
QY 61 QVBODEEDELTKYCAKHVIMLFVPTLCMVVAVTIKSVSEFYRKDQQLYTPETE 120  
Db 61 QVBODEEDELTKYCAKHVIMLFVPTLCMVVAVTIKSVSEFYRKDQQLYTPETE 120  
QY 121 DTEVGGKALHSLTAAIMISYIVYMTLLVLYLKRYCYKIHAMLTSSLLFFESFI 180  
Db 121 DTEVGGKALHSLTAAIMISYIVYMTLLVLYLKRYCYKIHAMLTSSLLFFESFI 180  
QY 181 YLGEVFKTYNAVDYITVALLIMNLGVGMISIMHKGPLRLQOAYLIMISALMALVFRTKY 240  
Db 181 YLGEVFKTYNAVDYITVALLIMNLGVGMISIMHKGPLRLQOAYLIMISALMALVFRTKY 240  
QY 241 LPEWTAMLLIAVISYVDLVAVLCPKGPLRMLVETAQERNETLFPALITYSTMVNLVNNAE 300  
Db 241 LPEWTAMLLIAVISYVDLVAVLCPKGPLRMLVETAQERNETLFPALITYSTMVNLVNNAE 300  
QY 301 GDPAQRVRSKNSKYNAESTRESODTYVAENDDGFSEEMEAQRDSDLGPHRSTPESRAA 360  
Db 301 GDPAQRVRSKNSKYNAESTRESODTYVAENDDGFSEEMEAQRDSDLGPHRSTPESRAA 360  
QY 361 VOELSSSILAGEDPEERGVKLGDFIFYSVLGKASATASGDMNTTACFVAILIGLCL 420  
Db 361 VOELSSSILAGEDPEERGVKLGDFIFYSVLGKASATASGDMNTTACFVAILIGLCL 420  
QY 421 TLLLLAIFKKALPALPISITFGLVFYFATDYLVOFMDQLAFHOFYI 467  
Db 421 TLLLLAIFKKALPALPISITFGLVFYFATDYLVOFMDQLAFHOFYI 467

RESULT 11  
US-08-670-964-2  
: Sequence 2, Application US/08670964  
: Patent No. 6010874  
: GENERAL INFORMATION:  
: APPLICANT: Hardy, John A.  
: TITLE OF INVENTION: EARLY ONSET ALZHEIMER'S DISEASE  
: NUMBER OF SEQUENCES: 4  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: SmithKline Beecham Corporation  
: STREET: 709 Swedeland Road - UW2220; P.O. Box 15  
: CITY: Philadelphia  
: STATE: PA  
: COUNTRY: USA  
: ZIP: 19406  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Diskette  
: OPERATING SYSTEM: IBM Compatible  
: SOFTWARE: FASTSEQ for Windows Version 2.0  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/670,964  
: FILING DATE: 26-JUN-1996  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: 60/001,142  
: FILING DATE: 13-JUL-1995  
: APPLICATION NUMBER: 60/001,501  
: FILING DATE: 18-JUL-1995  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Han, William T  
: REGISTRATION NUMBER: 34,344  
: REFERENCE/DOCKET NUMBER: P50358

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-5219  
TELEFAX: 610-270-5090

## TELEX:

INFORMATION FOR SEQ ID NO: 2:

## SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-670-964-2

## Query Match

Best Local Similarity 78.4%; Score 366; DB 3; Length 467;

Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MTELPAPLSYFQNAQMSDNHLSNTVRSQNDNREROHNDRLSGHPEPLSNRPGQNSR 60
D 1 MTELPAPLSYFQNAQMSDNHLSNTVRSQNDNREROHNDRLSGHPEPLSNRPGQNSR 60
QY 61 QVVEODEEDELTLKYGAKHVMLEVPVTLCAVNVVATIKVSFYTRKDGQLIYPTPE 120
D 61 QVVEODEEDELTLKYGAKHVMLEVPVTLCAVNVVATIKVSFYTRKDGQLIYPTPE 120
QY 121 DTEVGGALHSILNAIMSVIVMTLLVLYKRCYKVIHAWLIISLLFFESFI 180
D 121 DTEVGGALHSILNAIMSVIVMTLLVLYKRCYKVIHAWLIISLLFFESFI 180
QY 181 YLGEVFTYVAVDYITVALLINMFGVGMISIMKGPLRLOQAYLIMISALMALVFIKY 240
D 181 YLGEVFTYVAVDYITVALLINMFGVGMISIMKGPLRLOQAYLIMISALMALVFIKY 240
QY 241 LPEWTAMLLIAVISYVDLVAVLCPKGPLRMLVETAOERNETLFPALISSTWVLVMAE 300
D 241 LPEWTAMLLIAVISYVDLVAVLCPKGPLRMLVETAOERNETLFPALISSTWVLVMAE 300
QY 301 GDPAQRVSKNSKYNAESTERESQDTVAENDDGFSEWEAQRDHLGPHRSTPESRAA 360
D 301 GDPAQRVSKNSKYNAESTERESQDTVAENDDGFSEWEAQRDHLGPHRSTPESRAA 360
QY 361 VOELSSILAGEDPEERGVKLGDFEFYSVLGKASATASGDMNTTIACFVALILGLCL 420
D 361 VOELSSILAGEDPEERGVKLGDFEFYSVLGKASATASGDMNTTIACFVALILGLCL 420
QY 421 TLLLAIFKKALPALPISITFGVLFYFATDYLVQPFMDQLAFHQFYI 467
D 421 TLLLAIFKKALPALPISITFGVLFYFATDYLVQPFMDQLAFHQFYI 467
```

## RESULT 12

US-08-888-077A-2

Sequence 2, Application US/0888077A

Patent No. 6020143

## GENERAL INFORMATION:

APPLICANT: ST. GEORGE-HYSLOP, PETER H

APPLICANT: ROMMENS, JOHANNA M

APPLICANT: FRASER, PAUL E

TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED

TITLE OF INVENTION: TO ALZHEIMER'S DISEASE AND USES THEREFOR.

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:

ADDRESSEE: LERNER, DAVID, LITTENBERG, KRONHOLZ &amp; MENTLIK

STREET: 600 SOUTH AVENUE WEST

CITY: WESTFIELD

STATE: NJ

COUNTRY: USA

ZIP: 07090-1497

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/888,077A

FILING DATE: 03-JUL-1997

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/592,541

FILING DATE: 26-JAN-1996

ATTORNEY/AGENT INFORMATION:

NAME: PALISI, THOMAS M

REGISTRATION NUMBER: 36,629

REFERENCE/DOCKET NUMBER: SCHERING 3.0-017 CIP CIP CIP IV

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 654-5000

TELEFAX: (908) 654-7866

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-888-077A-2

## Query Match

Best Local Similarity 78.4%; Score 366; DB 3; Length 467;

Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MTELPAPLSYFQNAQMSDNHLSNTVRSQNDNREROHNDRLSGHPEPLSNRPGQNSR 60
D 1 MTELPAPLSYFQNAQMSDNHLSNTVRSQNDNREROHNDRLSGHPEPLSNRPGQNSR 60
QY 61 QVVEODEEDELTLKYGAKHVMLEVPVTLCAVNVVATIKVSFYTRKDGQLIYPTPE 120
D 61 QVVEODEEDELTLKYGAKHVMLEVPVTLCAVNVVATIKVSFYTRKDGQLIYPTPE 120
QY 121 DTEVGGALHSILNAIMSVIVMTLLVLYKRCYKVIHAWLIISLLFFESFI 180
D 121 DTEVGGALHSILNAIMSVIVMTLLVLYKRCYKVIHAWLIISLLFFESFI 180
QY 181 YLGEVFTYVAVDYITVALLINMFGVGMISIMKGPLRLOQAYLIMISALMALVFIKY 240
D 181 YLGEVFTYVAVDYITVALLINMFGVGMISIMKGPLRLOQAYLIMISALMALVFIKY 240
QY 241 LPEWTAMLLIAVISYVDLVAVLCPKGPLRMLVETAOERNETLFPALISSTWVLVMAE 300
D 241 LPEWTAMLLIAVISYVDLVAVLCPKGPLRMLVETAOERNETLFPALISSTWVLVMAE 300
QY 301 GDPAQRVSKNSKYNAESTERESQDTVAENDDGFSEWEAQRDHLGPHRSTPESRAA 360
D 301 GDPAQRVSKNSKYNAESTERESQDTVAENDDGFSEWEAQRDHLGPHRSTPESRAA 360
QY 361 VOELSSILAGEDPEERGVKLGDFEFYSVLGKASATASGDMNTTIACFVALILGLCL 420
D 361 VOELSSILAGEDPEERGVKLGDFEFYSVLGKASATASGDMNTTIACFVALILGLCL 420
QY 421 TLLLAIFKKALPALPISITFGVLFYFATDYLVQPFMDQLAFHQFYI 467
D 421 TLLLAIFKKALPALPISITFGVLFYFATDYLVQPFMDQLAFHQFYI 467
```

## RESULT 13

US-09-124-698-134

Sequence 134, Application US/09124698

Patent No. 6117978

## GENERAL INFORMATION:

APPLICANT: ST. GEORGE-HYSLOP, PETER H

APPLICANT: ROMMENS, JOHANNA M

APPLICANT: FRASER, PAUL E

TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED

TITLE OF INVENTION: TO ALZHEIMER'S DISEASE

NUMBER OF SEQUENCES: 183

CORRESPONDENCE ADDRESS:

ADDRESSEE: TESTA, HURWITZ &amp; THIBEAULT

STREET: High Street Tower - 125 High Street



CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/124,698  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/532,541  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 134:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-124-698-134

Query Match 78.4%; Score 366; DB 3; Length 467;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPELPALSYFQNAQSEDNHLSNTYRSONDNREOHNDRLSLGHEPLNSGRPGNSR 60  
DB 1 MPELPALSYFQNAQSEDNHLSNTYRSONDNREOHNDRLSLGHEPLNSGRPGNSR 60  
QY 61 OVEODEEDELTLKYGAKHVIMLFVPTLCWVVVATIKSVFTRKDGOLITPPT 120  
DB 61 OVEODEEDELTLKYGAKHVIMLFVPTLCWVVVATIKSVFTRKDGOLITPPT 120  
QY 121 DPEVQORALHSLNAINMISVIVMTILLVLYKRYKVIHAWLISSLLFFFSFT 180  
DB 121 DPEVQORALHSLNAINMISVIVMTILLVLYKRYKVIHAWLISSLLFFFSFT 180  
QY 181 YGGEVKTNNVADVITVALLLIMNFGVGMISIMHGKPLRLOQAYLIMISALMALVFIRY 240  
DB 181 YGGEVKTNNVADVITVALLLIMNFGVGMISIMHGKPLRLOQAYLIMISALMALVFIRY 240  
QY 241 LPEWTAMLILAVISYDVAVLCPRKPLMLVETAERNETLFPALISSTWMLVNM 300  
DB 241 LPEWTAMLILAVISYDVAVLCPRKPLMLVETAERNETLFPALISSTWMLVNM 300  
QY 301 GDEPAORRVSKSKYNAESTERESQDTVAENDGCFSEMEARORDSHLCPHRSTPESRAA 360  
DB 301 GDEPAORRVSKSKYNAESTERESQDTVAENDGCFSEMEARORDSHLCPHRSTPESRAA 360  
QY 361 VOELSSIIAGDEPBERGVKLGIDFTFYSVLVGKASATASGDMNTTICFAVAILGLCL 420  
DB 361 VOELSSIIAGDEPBERGVKLGIDFTFYSVLVGKASATASGDMNTTICFAVAILGLCL 420  
QY 421 TLLLLAIFKKALPALPISITFGLVFEATDYLVOPMDOLAFHOFT 467  
DB 421 TLLLLAIFKKALPALPISITFGLVFEATDYLVOPMDOLAFHOFT 467

RESULT 14  
US-09-127-480-134  
Sequence 134, Application US/09127480  
Patent No. 6194153  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H

APPLICANT: ROMENS, JOHANNA M  
APPLICANT: FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TO ALZHEIMER'S DISEASE  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
STREET: High Street Tower - 125 High Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/127,480  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/532,541  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 134:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-127-480-134

Query Match 78.4%; Score 366; DB 4; Length 467;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPELPALSYFQNAQSEDNHLSNTYRSONDNREOHNDRLSLGHEPLNSGRPGNSR 60  
DB 1 MPELPALSYFQNAQSEDNHLSNTYRSONDNREOHNDRLSLGHEPLNSGRPGNSR 60  
QY 61 OVEODEEDELTLKYGAKHVIMLFVPTLCWVVVATIKSVFTRKDGOLITPPT 120  
DB 61 OVEODEEDELTLKYGAKHVIMLFVPTLCWVVVATIKSVFTRKDGOLITPPT 120  
QY 121 DPEVQORALHSLNAINMISVIVMTILLVLYKRYKVIHAWLISSLLFFFSFT 180  
DB 121 DPEVQORALHSLNAINMISVIVMTILLVLYKRYKVIHAWLISSLLFFFSFT 180  
QY 181 YGGEVKTNNVADVITVALLLIMNFGVGMISIMHGKPLRLOQAYLIMISALMALVFIRY 240  
DB 181 YGGEVKTNNVADVITVALLLIMNFGVGMISIMHGKPLRLOQAYLIMISALMALVFIRY 240  
QY 241 LPEWTAMLILAVISYDVAVLCPRKPLMLVETAERNETLFPALISSTWMLVNM 300  
DB 241 LPEWTAMLILAVISYDVAVLCPRKPLMLVETAERNETLFPALISSTWMLVNM 300  
QY 301 GDEPAORRVSKSKYNAESTERESQDTVAENDGCFSEMEARORDSHLCPHRSTPESRAA 360  
DB 301 GDEPAORRVSKSKYNAESTERESQDTVAENDGCFSEMEARORDSHLCPHRSTPESRAA 360  
QY 361 VOELSSIIAGDEPBERGVKLGIDFTFYSVLVGKASATASGDMNTTICFAVAILGLCL 420  
DB 361 VOELSSIIAGDEPBERGVKLGIDFTFYSVLVGKASATASGDMNTTICFAVAILGLCL 420  
QY 421 TLLLLAIFKKALPALPISITFGLVFEATDYLVOPMDOLAFHOFT 467  
DB 421 TLLLLAIFKKALPALPISITFGLVFEATDYLVOPMDOLAFHOFT 467

RESULT 15  
US-08-496-841C-134  
; Sequence 134, Application US/08496841C  
; Patent No. 6210919  
; GENERAL INFORMATION:  
; APPLICANT: ST. GEORGE-HYSLOP, PETER H  
; ROMMENS, JOHANNA M  
; FRASER, PAUL E  
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
; TO ALZHEIMER'S DISEASE  
; NUMBER OF SEQUENCES: 175  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Darryl & Darby, PC  
; STREET: 805 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/496,841C  
; FILING DATE: 28-JUN-1995  
; CLASSIFICATION: <unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Paul F. Fehner, Ph.D.  
; REGISTRATION NUMBER: 35,135  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 527-7700  
; TELEFAX: (212) 753-6237  
; INFORMATION FOR SEQ ID NO: 134:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 467 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 134:  
US-08-496-841C-134  
Query Match 78.4%; Score 366; DB 4; Length 467;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MTELPAPLSYFONAOQSEDNHLSNTVRSQNDNREROEHNDRRSLGHPPLSNGRQGSNR 60  
DB 1 MTELPAPLSYFONAOQSEDNHLSNTVRSQNDNREROEHNDRRSLGHPPLSNGRQGSNR 60  
QY 61 QVVEDEDEDELTLKYGAKHYIMLFVPTLCMVVVVATIKSVSYTRKDGQLITTPETE 120  
DB 61 QVVEDEDEDELTLKYGAKHYIMLFVPTLCMVVVVATIKSVSYTRKDGQLITTPETE 120  
QY 121 DTEYGORALHSILNAAIMISYIVMTLLVLYKYRCYKVIHAWLISSLLFFFSFI 180  
DB 121 DTEYGORALHSILNAAIMISYIVMTLLVLYKYRCYKVIHAWLISSLLFFFSFI 180  
QY 181 YLGEVFKTYNNAVDTYVALLIMNGVGMISIHMKGPLRLQOAVLIMISALMALVFITY 240  
DB 181 YLGEVFKTYNNAVDTYVALLIMNGVGMISIHMKGPLRLQOAVLIMISALMALVFITY 240  
QY 241 LPEWTAMILLAVISYDLYAVLCPPKPLRLMVLVETAOERNETLFPALISSTWMLVNNAE 300  
DB 241 LPEWTAMILLAVISYDLYAVLCPPKPLRLMVLVETAOERNETLFPALISSTWMLVNNAE 300  
QY 301 GDPFAORRVSKSKYNASTERESODTVAENDDGFSEMEQORDSHLGPHRSTPESRAA 360  
DB 301 GDPFAORRVSKSKYNASTERESODTVAENDDGFSEMEQORDSHLGPHRSTPESRAA 360

QY 361 VOELSSSILAGDEPBERGVKLGIDFIYSVLVGKASATASGDWMTTACFAAILIGLCL 420  
DB 361 VOELSSSILAGDEPBERGVKLGIDFIYSVLVGKASATASGDWMTTACFAAILIGLCL 420  
QY 421 TLLILAIFKKALPALPISITTFGLVFFATDYLVQPMQDLAFHOEYI 467  
DB 421 TLLILAIFKKALPALPISITTFGLVFFATDYLVQPMQDLAFHOEYI 467  
RESULT 16  
US-08-832-867-3  
; Sequence 3, Application US/08832867C  
; Patent No. 6376239  
; GENERAL INFORMATION:  
; APPLICANT: BAUMEISTER, Ralf  
; TITLE OF INVENTION: DNA MOLECULES COMPRISING A PROMOTER CAPABLE OF  
; TITLE OF INVENTION: CONFERRING EXPRESSION OF A HETEROLOGOUS DNA SEQUENCE IN  
; TITLE OF INVENTION: ALL NEURAL CELLS AT ALL STAGES OF DEVELOPMENT IN C.  
; FILE REFERENCE: 674503-2004  
; CURRENT APPLICATION NUMBER: US/08/832,867C  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 467  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
US-08-832-867-3  
Query Match 78.4%; Score 366; DB 4; Length 467;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MTELPAPLSYFONAOQSEDNHLSNTVRSQNDNREROEHNDRRSLGHPPLSNGRQGSNR 60  
DB 1 MTELPAPLSYFONAOQSEDNHLSNTVRSQNDNREROEHNDRRSLGHPPLSNGRQGSNR 60  
QY 61 QVVEDEDEDELTLKYGAKHYIMLFVPTLCMVVVVATIKSVSYTRKDGQLITTPETE 120  
DB 61 QVVEDEDEDELTLKYGAKHYIMLFVPTLCMVVVVATIKSVSYTRKDGQLITTPETE 120  
QY 121 DTEYGORALHSILNAAIMISYIVMTLLVLYKYRCYKVIHAWLISSLLFFFSFI 180  
DB 121 DTEYGORALHSILNAAIMISYIVMTLLVLYKYRCYKVIHAWLISSLLFFFSFI 180  
QY 181 YLGEVFKTYNNAVDTYVALLIMNGVGMISIHMKGPLRLQOAVLIMISALMALVFITY 240  
DB 181 YLGEVFKTYNNAVDTYVALLIMNGVGMISIHMKGPLRLQOAVLIMISALMALVFITY 240  
QY 241 LPEWTAMILLAVISYDLYAVLCPPKPLRLMVLVETAOERNETLFPALISSTWMLVNNAE 300  
DB 241 LPEWTAMILLAVISYDLYAVLCPPKPLRLMVLVETAOERNETLFPALISSTWMLVNNAE 300  
QY 301 GDPFAORRVSKSKYNASTERESODTVAENDDGFSEMEQORDSHLGPHRSTPESRAA 360  
DB 301 GDPFAORRVSKSKYNASTERESODTVAENDDGFSEMEQORDSHLGPHRSTPESRAA 360  
QY 361 VOELSSSILAGDEPBERGVKLGIDFIYSVLVGKASATASGDWMTTACFAAILIGLCL 420  
DB 361 VOELSSSILAGDEPBERGVKLGIDFIYSVLVGKASATASGDWMTTACFAAILIGLCL 420  
QY 421 TLLILAIFKKALPALPISITTFGLVFFATDYLVQPMQDLAFHOEYI 467  
DB 421 TLLILAIFKKALPALPISITTFGLVFFATDYLVQPMQDLAFHOEYI 467

RESULT 17  
US-09-227-725A-1  
; Sequence 1, Application US/09227725A  
; Patent No. 6383758  
; GENERAL INFORMATION:  
; APPLICANT: St. George-Hyslop, Peter H.

APPLICANT: Rommens, Johanna  
APPLICANT: Fraser, Paul E.  
TITLE OF INVENTION: Alzheimer's Related Proteins and Methods  
TITLE OF INVENTION: of use  
FILE REFERENCE: 1034/1810-US1  
CURRENT APPLICATION NUMBER: US/09/227,725A  
CURRENT FILING DATE: 1999-01-08  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO: 1  
LENGTH: 467  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-09-227-725A-1

Query Match 78.4%; Score 366; DB 4; Length 467;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTELPAFLSYFONAKSEDNHLSNTVRSONDNREOEHNDRSLGHPPELNGRPOGNSR 60  
DB 1 MTELPAFLSYFONAKSEDNHLSNTVRSONDNREOEHNDRSLGHPPELNGRPOGNSR 60  
QY 61 QVEQDEEDELTLKYGAKHVMLEFPVTLQVNVVATIKSVFTRKDGOLITPPE 120  
DB 61 QVEQDEEDELTLKYGAKHVMLEFPVTLQVNVVATIKSVFTRKDGOLITPPE 120  
QY 121 DPEYVQGRALHSLNAAIMISVIVMTLLVLYKYRCYKVIHAWLIISLLFFESFI 180  
DB 121 DPEYVQGRALHSLNAAIMISVIVMTLLVLYKYRCYKVIHAWLIISLLFFESFI 180  
QY 181 YLGEVETKYNVAVDYITVALLINNLGVGMISIMHKGPRLOQAYLIMISALMALVFITY 240  
DB 181 YLGEVETKYNVAVDYITVALLINNLGVGMISIMHKGPRLOQAYLIMISALMALVFITY 240  
QY 241 LPEWTAMLLIAVSVYDLVAVLCPKGPLRMLVETAOERNETLFPALISSTWMLVMAE 300  
DB 241 LPEWTAMLLIAVSVYDLVAVLCPKGPLRMLVETAOERNETLFPALISSTWMLVMAE 300  
QY 301 GDEPAORRVSKSKYNAESTERESODTVAENDGGSSEMEAOARDSHLGPHRSTPESRAA 360  
DB 301 GDEPAORRVSKSKYNAESTERESODTVAENDGGSSEMEAOARDSHLGPHRSTPESRAA 360  
QY 361 VOELSSILAGEDPEERGVKLGIDFTFYSVLGKASATASGDMNTTIACFVAIILGLCL 420  
DB 361 VOELSSILAGEDPEERGVKLGIDFTFYSVLGKASATASGDMNTTIACFVAIILGLCL 420  
QY 421 TLLLLAIFKKALPALPISITFGLVFATDYLVOPEMDOLAFHOFTY 467  
DB 421 TLLLLAIFKKALPALPISITFGLVFATDYLVOPEMDOLAFHOFTY 467

RESULT 18  
US-09-124-523-134  
Sequence 134, Application US/09124523  
Patent No. 6395960

GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: ROMMENS, JOHANNA M  
APPLICANT: FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HUMMIZ & THIBEAULT  
STREET: High Street Tower - 125 High Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/124,523  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,541  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 134:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-124-523-134

Query Match 78.4%; Score 366; DB 4; Length 467;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTELPAFLSYFONAKSEDNHLSNTVRSONDNREOEHNDRSLGHPPELNGRPOGNSR 60  
DB 1 MTELPAFLSYFONAKSEDNHLSNTVRSONDNREOEHNDRSLGHPPELNGRPOGNSR 60  
QY 61 QVEQDEEDELTLKYGAKHVMLEFPVTLQVNVVATIKSVFTRKDGOLITPPE 120  
DB 61 QVEQDEEDELTLKYGAKHVMLEFPVTLQVNVVATIKSVFTRKDGOLITPPE 120  
QY 121 DPEYVQGRALHSLNAAIMISVIVMTLLVLYKYRCYKVIHAWLIISLLFFESFI 180  
DB 121 DPEYVQGRALHSLNAAIMISVIVMTLLVLYKYRCYKVIHAWLIISLLFFESFI 180  
QY 181 YLGEVETKYNVAVDYITVALLINNLGVGMISIMHKGPRLOQAYLIMISALMALVFITY 240  
DB 181 YLGEVETKYNVAVDYITVALLINNLGVGMISIMHKGPRLOQAYLIMISALMALVFITY 240  
QY 241 LPEWTAMLLIAVSVYDLVAVLCPKGPLRMLVETAOERNETLFPALISSTWMLVMAE 300  
DB 241 LPEWTAMLLIAVSVYDLVAVLCPKGPLRMLVETAOERNETLFPALISSTWMLVMAE 300  
QY 301 GDEPAORRVSKSKYNAESTERESODTVAENDGGSSEMEAOARDSHLGPHRSTPESRAA 360  
DB 301 GDEPAORRVSKSKYNAESTERESODTVAENDGGSSEMEAOARDSHLGPHRSTPESRAA 360  
QY 361 VOELSSILAGEDPEERGVKLGIDFTFYSVLGKASATASGDMNTTIACFVAIILGLCL 420  
DB 361 VOELSSILAGEDPEERGVKLGIDFTFYSVLGKASATASGDMNTTIACFVAIILGLCL 420  
QY 421 TLLLLAIFKKALPALPISITFGLVFATDYLVOPEMDOLAFHOFTY 467  
DB 421 TLLLLAIFKKALPALPISITFGLVFATDYLVOPEMDOLAFHOFTY 467

RESULT 19  
US-09-375-318-3  
Sequence 3, Application US/09375318  
Patent No. 6468791

GENERAL INFORMATION:  
APPLICANT: Tanzi, Rudolph E.  
Schellenberg, Gerard D.  
Masco, Wilma

Levy-Lahad, Ephrat  
Bitd, Thomas D.  
Galas, David J.  
TITLE OF INVENTION: CHROMOSOME 1 GENE AND GENE PRODUCTS RELATED TO  
ALZHEIMER'S DISEASE

NUMBER OF SEQUENCES: 88  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SEED and BEERY LLP  
 STREET: 701 Fifth Ave, Suite 6300  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: USA  
 ZIP: 98104-7092  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/375,318  
 FILING DATE: 16-Aug-1999  
 CLASSIFICATION: <unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Verna, James M.  
 REGISTRATION NUMBER: 33,287  
 REFERENCE/DOCKET NUMBER: 920010.571C1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 467 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: <unknown>  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
 US-09-375-318-3  
 Query Match 78.4%; Score 366; DB 4; Length 467;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MTELPAPLSYFQNAQMSQEDNHLSTNVRSONDNREORHNDRLSLGHPLELSNGRQGNR 60  
 DB 1 MTELPAPLSYFQNAQMSQEDNHLSTNVRSONDNREORHNDRLSLGHPLELSNGRQGNR 60  
 QY 61 QVVEODEDEDELTLTKYGAHVIMLFVPTLCMVVVVATIKSVFTYTRKQGLIYTPTE 120  
 DB 61 QVVEODEDEDELTLTKYGAHVIMLFVPTLCMVVVVATIKSVFTYTRKQGLIYTPTE 120  
 QY 121 DFEYFGORLHSLNAAIMISIVYVMTLLVLYKYRCYKVIHAWLIISLLLEFESPI 180  
 DB 121 DFEYFGORLHSLNAAIMISIVYVMTLLVLYKYRCYKVIHAWLIISLLLEFESPI 180  
 QY 181 YLGEVFKTYNVAVDYITVALLIMNLGVGMISIHMKGPLRLQOAVLIMISALMALVFIRY 240  
 DB 181 YLGEVFKTYNVAVDYITVALLIMNLGVGMISIHMKGPLRLQOAVLIMISALMALVFIRY 240  
 QY 241 LPEWTAWLLAVISYVDLAVLCRPGPLMLVETAOERNETLFPALIISSSTWMLVNNAE 300  
 DB 241 LPEWTAWLLAVISYVDLAVLCRPGPLMLVETAOERNETLFPALIISSSTWMLVNNAE 300  
 QY 301 GDPPEORRRSKSKYNAESTERESODTVAENDGGFSEEWPAQRDSHLGPSTESRA 360  
 DB 301 GDPPEORRRSKSKYNAESTERESODTVAENDGGFSEEWPAQRDSHLGPSTESRA 360  
 QY 361 VOELSSSILAGDPERGVKLGDFIFYSVLVKASATASGDWMTTACFAVAILIGLCL 420  
 DB 361 VOELSSSILAGDPERGVKLGDFIFYSVLVKASATASGDWMTTACFAVAILIGLCL 420  
 QY 421 TLLLAIFKKALPALPISTITGVLVFPATDYLVOPMDQLAFHOFYI 467  
 DB 421 TLLLAIFKKALPALPISTITGVLVFPATDYLVOPMDQLAFHOFYI 467  
 RESULT 20  
 US-08-670-964-4  
 Sequence 4, Application US/08670964

Patent No. 6010874  
 GENERAL INFORMATION:  
 APPLICANT: Hardy, John A.  
 TITLE OF INVENTION: EARLY ONSET ALZHEIMER'S DISEASE  
 TITLE OF INVENTION: GENE AND GENE PRODUCTS  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SmithKline Beecham Corporation  
 STREET: 709 Swedeland Road - UW2220; P.O. Box 15  
 CITY: Philadelphia  
 STATE: PA  
 COUNTRY: USA  
 ZIP: 19406  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/670,964  
 FILING DATE: 26-JUN-1996  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/001,142  
 FILING DATE: 3-JUL-1995  
 APPLICATION NUMBER: 60/001,501  
 FILING DATE: 18-JUL-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Han, William T.  
 REGISTRATION NUMBER: 34,344  
 REFERENCE/DOCKET NUMBER: P50358  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 610-270-5219  
 TELEFAX: 610-270-5090  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 463 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-670-964-4  
 Query Match 72.2%; Score 337; DB 3; Length 463;  
 Best Local Similarity 99.8%; Pred. No. 5.8e-313;  
 Matches 437; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 30 NDNREORHNDRLSLGHPLELSNGRQGNRQVYQDEDEDELTLKYGAHVIMLFV 89  
 DB 26 NDNREORHNDRLSLGHPLELSNGRQGNRQVYQDEDEDELTLKYGAHVIMLFV 85  
 QY 90 TLCMVVVVATIKSVFTYTRKQGLIYTPTEDETFVGOALHSIINAAIMISIVYVMTL 149  
 DB 86 TLCMVVVVATIKSVFTYTRKQGLIYTPTEDETFVGOALHSIINAAIMISIVYVMTL 145  
 QY 150 LVLYKYRCYKVIHAWLIISLLLEFESFYTLGEVFKTYNVAVDYITVALLIMNLGVG 209  
 DB 146 LVLYKYRCYKVIHAWLIISLLLEFESFYTLGEVFKTYNVAVDYITVALLIMNLGVG 205  
 QY 210 MSHMKGPLRLQOAVLIMISALMALVFIRYKLPENTAMLLAVISYVDLAVLCPKGPLR 269  
 DB 206 MSHMKGPLRLQOAVLIMISALMALVFIRYKLPENTAMLLAVISYVDLAVLCPKGPLR 265  
 QY 270 MLVETAOERNETLFPALIISSSTWMLVNNAEQDPPEORRRSKSKYNAESTERESODTVA 329  
 DB 266 MLVETAOERNETLFPALIISSSTWMLVNNAEQDPPEORRRSKSKYNAESTERESODTVA 325  
 QY 330 ENDDGFESEEWPAQRDSHLGPSTESRAVVOELSSSILAGDPERGVKLGDFIFY 389  
 DB 326 ENDDGFESEEWPAQRDSHLGPSTESRAVVOELSSSILAGDPERGVKLGDFIFY 385  
 QY 390 SVLVKASATASGDWMTTACFAVAILIGLCLTLLLAIFKKALPALPISTITGVLVFPAT 449

Db 386 SYLVAKASATASGDMNTTACFAVAILIGLCTLLLAIFKKALPALPISITFGLVFAT 445  
Oy 450 DYLVOPMDQLAHOFYI 467  
Db 446 DTLVOPMDQLAHOFYI 463

## RESULT 21

US-08-875-972-4  
Sequence 4, Application US/08875972  
Patent No. 5985564

GENERAL INFORMATION:  
APPLICANT: Huntington Potter and Jinhue Li  
TITLE OF INVENTION: ASSAY FOR IDENTIFYING GENES CAUSING  
TITLE OF INVENTION: CHROMOSOME NON-DISJUNCTION  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.  
STREET: Two Millitia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173-4799  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/875,972  
FILING DATE: 08-AUG-97  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/002,448  
FILING DATE: 16-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan Esq., Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: HU95-03PA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781) 861-6240  
TELEFAX: (781) 861-9540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 407 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-875-972-4

Query Match 65.5%; Score 306; DB 2; Length 407;

Best Local Similarity 99.8%; Pred. No. 1.8e-283;

Matches 406; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 61 OVEDEDEDEDELTLKYGKHHVIMLFVPTLCMVVVVATIKSVSYTRDGLIYTPETE 120  
Db 1 OVEDEDEDEDELTLKYGKHHVIMLFVPTLCMVVVVATIKSVSYTRDGLIYTPETE 60  
Oy 121 DTEVGORALHSILNAAMISIVVMTLLVLYKRCYKVIHAWLIISLLLEFFSFI 180  
Db 61 DTEVGORALHSILNAAMISIVVMTLLVLYKRCYKVIHAWLIISLLLEFFSFI 120  
Oy 181 YLGEVFKTYNNAVVDITVALLIWNLCVGMISIHMKGPLRLOQAVLIMISALMALVFIRY 240  
Db 121 YLGEVFKTYNNAVVDITVALLIWNLCVGMISIHMKGPLRLOQAVLIMISALMALVFIRY 180  
Oy 241 LPENTAMILAVISYVDLAVAVLCPRGPLMLVETAORNETLFPALITSSWTWVLYNMAE 300  
Db 181 LPENTAMILAVISYVDLAVAVLCPRGPLMLVETAORNETLFPALITSSWTWVLYNMAE 240  
Oy 301 GDEAQRVSKNSKYNAESTERESODTVAENDGGFSEMEQRSHLGPHRSTPESRAA 360  
Db 181 GDEAQRVSKNSKYNAESTERESODTVAENDGGFSEMEQRSHLGPHRSTPESRAA 360

Db 241 GDEAQRVSKNSKYNAESTERESODTVAENDGGFSEMEQRSHLGPHRSTPESRAA 300  
Oy 361 VOELSSSILAGDEPERGVKLGIDFIFSVYVGRASATASGDMNTTACFAVAILIGL 420  
Db 301 VOELSSSILAGDEPERGVKLGIDFIFSVYVGRASATASGDMNTTACFAVAILIGL 360  
Oy 421 TLLLAIFKKALPALPISITFGLVFATDYLVOFMDLAHOFYI 467  
Db 361 TLLLAIFKKALPALPISITFGLVFATDYLVOFMDLAHOFYI 407

## RESULT 22

US-08-706-344C-2  
Sequence 2, Application US/08706344C  
Patent No. 6248555

GENERAL INFORMATION:  
APPLICANT: TANZI, RUDOLPH  
APPLICANT: WASCO, WILMA  
TITLE OF INVENTION: Genetic Alterations Related To Familial  
TITLE OF INVENTION: Alzheimer's Disease  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVENUE, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (Epo)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/706,344C  
FILING DATE: 30-AUG-1996  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/003,054  
FILING DATE: 31-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: KIM, JUDITH U.  
REGISTRATION NUMBER: 40,679  
REFERENCE/DOCKET NUMBER: 0609.4180001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-706-344C-2

Query Match 56.7%; Score 265; DB 4; Length 467;

Best Local Similarity 99.6%; Pred. No. 2.5e-244;

Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 MTELPAPLISYFQNAQSEDNHLSNTVRSQNDNREROHNDRLSLGHPPLSGRPGQNSR 60  
Db 1 MTELPAPLISYFQNAQSEDNHLSNTVRSQNDNREROHNDRLSLGHPPLSGRPGQNSR 60  
Oy 61 OVEDEDEDEDELTLKYGKHHVIMLFVPTLCMVVVVATIKSVSYTRDGLIYTPETE 120  
Db 61 OVEDEDEDEDELTLKYGKHHVIMLFVPTLCMVVVVATIKSVSYTRDGLIYTPETE 120  
Oy 121 DTEVGORALHSILNAAMISIVVMTLLVLYKRCYKVIHAWLIISLLLEFFSFI 180  
Db 121 DTEVGORALHSILNAAMISIVVMTLLVLYKRCYKVIHAWLIISLLLEFFSFI 180  
Oy 181 YLGEVFKTYNNAVVDITVALLIWNLCVGMISIHMKGPLRLOQAVLIMISALMALVFIRY 240  
Db 181 YLGEVFKTYNNAVVDITVALLIWNLCVGMISIHMKGPLRLOQAVLIMISALMALVFIRY 240

Db 181 YLGEVFKTYNVAVDYITVALLIWNQVGMISIMKGPRLQOAVYIMISALMALVEIKY 240

Qy 241 LPEWTAMLLAVISYVDLVAVLCPKGPLMVLVETAOERNETLLFPALISSTWMLVNMME 300  
|||||

Db 241 LPEWTAMLLAVISYVDLVAVLCPKGPLMVLVETAOERNETLLFPALISSTWMLVNMME 300

Qy 301 GDBEORARYSKSKNAESTERESQDTVAENDDGGFSEMEAROSHLGPHSTPESRAA 360  
|||||

Db 301 GDBEORARYSKSKNAESTERESQDTVAENDDGGFSEMEAROSHLGPHSTPESRAA 360

Qy 361 VOELSSSILAGDEPBERGKLGIDPFIFYSVLVGNASATASGDMNTTACFVAILIGLCL 420  
|||||

Db 361 VOELSSSILAGDEPBERGKLGIDPFIFYSVLVGNASATASGDMNTTACFVAILIGLCL 420

Qy 421 TLLLAIFKKALPALPISITFGLVFFATDYLVPFMDLAFHOPIY 467  
|||||

Db 421 TLLLAIFKKALPALPISITFGLVFFATDYLVPFMDLAFHOPIY 467

RESULT 23  
US-08-706-344C-4  
Sequence 4, Application US/08706344C  
Patent No. 6248555  
GENERAL INFORMATION:  
APPLICANT: TANZI, RUDOLPH  
APPLICANT: WASCO, WILMA  
TITLE OF INVENTION: Genetic Alterations Related To Familial  
TITLE OF INVENTION: Alzheimer's Disease  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVENUE, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/706,344C  
FILING DATE: 30-AUG-1996  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/003,054  
FILING DATE: 31-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: KIM, JUDITH U.  
REGISTRATION NUMBER: 40,679  
REFERENCE/DOCKET NUMBER: 0609,4180001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-706-344C-4

Query Match 43.7%; Score 204; DB 4; Length 467;  
Best Local Similarity 100.0%; Pred. No. 3.4e-186;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTELPAPLSTYFQNMOMSEDNHLSNTVRSQNDNREROEHNDRRSLGHPPLSNGRPOGNSR 60  
|||||

Db 1 MTELPAPLSTYFQNMOMSEDNHLSNTVRSQNDNREROEHNDRRSLGHPPLSNGRPOGNSR 60

Qy 61 QVVEODEDEDELTLKYGAKHVMFLFVPTLCMVVVATISVSFYTRKDGQLLYTPETE 120  
|||||

Db 61 QVVEODEDEDELTLKYGAKHVMFLFVPTLCMVVVATISVSFYTRKDGQLLYTPETE 120

Qy 121 DTEYVGOARALHSILNAAMISVIYVMTLLVLYKRCYKVIHAMLISLLFFPSFI 180  
|||||

Db 121 DTEYVGOARALHSILNAAMISVIYVMTLLVLYKRCYKVIHAMLISLLFFPSFI 180

Qy 181 YLGEVFKTYNVAVDYITVALLIWN 204  
|||||

Db 181 YLGEVFKTYNVAVDYITVALLIWN 204

RESULT 24  
US-08-706-344C-28  
Sequence 28, Application US/08706344C  
Patent No. 6248555  
GENERAL INFORMATION:  
APPLICANT: TANZI, RUDOLPH  
APPLICANT: WASCO, WILMA  
TITLE OF INVENTION: Genetic Alterations Related To Familial  
TITLE OF INVENTION: Alzheimer's Disease  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVENUE, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/706,344C  
FILING DATE: 30-AUG-1996  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/003,054  
FILING DATE: 31-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: KIM, JUDITH U.  
REGISTRATION NUMBER: 40,679  
REFERENCE/DOCKET NUMBER: 0609,4180001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-706-344C-28

Query Match 43.7%; Score 204; DB 4; Length 467;  
Best Local Similarity 100.0%; Pred. No. 3.4e-186;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTELPAPLSTYFQNMOMSEDNHLSNTVRSQNDNREROEHNDRRSLGHPPLSNGRPOGNSR 60  
|||||

Db 1 MTELPAPLSTYFQNMOMSEDNHLSNTVRSQNDNREROEHNDRRSLGHPPLSNGRPOGNSR 60

Qy 61 QVVEODEDEDELTLKYGAKHVMFLFVPTLCMVVVATISVSFYTRKDGQLLYTPETE 120  
|||||

Db 61 QVVEODEDEDELTLKYGAKHVMFLFVPTLCMVVVATISVSFYTRKDGQLLYTPETE 120

Qy 121 DTEYVGOARALHSILNAAMISVIYVMTLLVLYKRCYKVIHAMLISLLFFPSFI 180  
|||||

Db 121 DTEYVGOARALHSILNAAMISVIYVMTLLVLYKRCYKVIHAMLISLLFFPSFI 180

Qy 181 YLGEVFKTYNVAVDYITVALLIWN 204  
|||||



```

STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/788,231A
FILING DATE: 24-JAN-1997
CLASSIFICATION: 42A
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/010,672
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 60-95
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: YES
US-08-788-231A-15

```

```

Query Match          40.5%; Score 189; DB 3; Length 462;
Best Local Similarity 100.0%; Pred. No. 6,6e-172;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 206 GVGCMISHWKGPLRLQOAYLIMISALMAVFYKLPWTAMLILAVISYVDLVAVLCPR 265
    |||||||
DB 201 GVGCMISHWKGPLRLQOAYLIMISALMAVFYKLPWTAMLILAVISYVDLVAVLCPR 260
    |||||||
QY 266 GPLMLVETAOERNEETLPALIVSSTWVWLVNMAEGDEAQRVSKSKYNAESTERESQ 325
    |||||||
DB 261 GPLMLVETAOERNEETLPALIVSSTWVWLVNMAEGDEAQRVSKSKYNAESTERESQ 320
    |||||||
QY 326 DTVANDDGGFSEMEARQDSHLGPRHSTPESRAAVOELSSILAGDEPBERGVKLGID 385
    |||||||
DB 321 DTVANDDGGFSEMEARQDSHLGPRHSTPESRAAVOELSSILAGDEPBERGVKLGID 380
    |||||||
QY 386 FIFYSVLVG 394
    |||||||
DB 381 FIFYSVLVG 389

```

```

RESULT 28
US-08-670-479-18
Sequence 18, Application US/08670479
Patent No. 5973133
GENERAL INFORMATION:
APPLICANT: Hardy, John A.
APPLICANT: Goate, Allison M.
TITLE OF INVENTION: MUTANT S182 GENES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESS: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

```

```

SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/670,479
FILING DATE: 26-JUN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,500
FILING DATE: 18-JUL-1996
APPLICATION NUMBER: 60/001,800
FILING DATE: 02-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Han, William T.
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: P50361
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5219
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
US-08-670-479-18

```

```

Query Match          40.3%; Score 188; DB 2; Length 463;
Best Local Similarity 100.0%; Pred. No. 5,9e-171;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 280 ETLFPALIVSSTWVWLVNMAEGDEAQRVSKSKYNAESTERESQDTVAENDDGGFSEE 339
    |||||||
DB 276 ETLFPALIVSSTWVWLVNMAEGDEAQRVSKSKYNAESTERESQDTVAENDDGGFSEE 335
    |||||||
QY 340 WEAQRDSHLGPRHSTPESRAAVOELSSILAGDEPBERGVKLGIDTFYSVLVGKASAT 399
    |||||||
DB 336 WEAQRDSHLGPRHSTPESRAAVOELSSILAGDEPBERGVKLGIDTFYSVLVGKASAT 395
    |||||||
QY 400 ASGDMNTTACFAVAILGLCTLLLLAIFKKALPALPISITFGIVFFATDYLQPPMDQ 459
    |||||||
DB 396 ASGDMNTTACFAVAILGLCTLLLLAIFKKALPALPISITFGIVFFATDYLQPPMDQ 455
    |||||||
QY 460 LAHQFYI 467
    |||||||
DB 456 LAHQFYI 463

```

```

RESULT 29
US-08-967-101-4
Sequence 4, Application US/08967101
Patent No. 5840540
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESS: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```



```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,101
; FILING DATE: 10-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-967-101-4

Query Match                20.8%; Score 97; DB 2; Length 467;
Best Local Similarity      100.0%; Pred. No. 3.2e-84;
Matches 97: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 213 IHMKGPLRQOAYLMISALMALVFIKYPENTAMILLAVISYDVLVAVLCRKGPLRMIV 272
Db 213 IHMKGPLRQOAYLMISALMALVFIKYPENTAMILLAVISYDVLVAVLCRKGPLRMIV 272

OY 273 ETAOERNETLPALITYSSMTWMLVNMAEGDPEAORRV 309
Db 273 ETAOERNETLPALITYSSMTWMLVNMAEGDPEAORRV 309

RESULT 30
US-08-592-541-4
; Sequence 4, Application US/08592541
; Patent No. 5986054
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,541
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-592-541-4

Query Match                20.8%; Score 97; DB 2; Length 467;
Best Local Similarity      100.0%; Pred. No. 3.2e-84;
Matches 97: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 213 IHMKGPLRQOAYLMISALMALVFIKYPENTAMILLAVISYDVLVAVLCRKGPLRMIV 272
Db 213 IHMKGPLRQOAYLMISALMALVFIKYPENTAMILLAVISYDVLVAVLCRKGPLRMIV 272

OY 273 ETAOERNETLPALITYSSMTWMLVNMAEGDPEAORRV 309
Db 273 ETAOERNETLPALITYSSMTWMLVNMAEGDPEAORRV 309

RESULT 31
US-08-888-077A-17
; Sequence 17, Application US/08888077A
; Patent No. 6020143
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE AND USES THEREFOR.
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LERNER, DAVID, LITTEMBERG, KRUMHOLZ & MENTLIK
; STREET: 600 SOUTH AVENUE WEST
; CITY: WESTFIELD
; STATE: NJ
; COUNTRY: USA
; ZIP: 07090-1497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/888,077A
; FILING DATE: 03-JUL-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,541
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: PALISI, THOMAS M
; REGISTRATION NUMBER: 36,629
; REFERENCE/DOCKET NUMBER: SCHERING 3.0-017 CIP CIP IV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 654-5000
; TELEFAX: (908) 654-7866
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-888-077A-17

Query Match                20.8%; Score 97; DB 3; Length 467;
Best Local Similarity      100.0%; Pred. No. 3.2e-84;
Matches 97: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 213 IHMKGPLRQOAYLMISALMALVFIKYPENTAMILLAVISYDVLVAVLCRKGPLRMIV 272
Db 213 IHMKGPLRQOAYLMISALMALVFIKYPENTAMILLAVISYDVLVAVLCRKGPLRMIV 272

OY 273 ETAOERNETLPALITYSSMTWMLVNMAEGDPEAORRV 309
Db 273 ETAOERNETLPALITYSSMTWMLVNMAEGDPEAORRV 309
```

RESULT 32  
US-09-124-698-4  
Sequence 4, Application US/09124698  
Patent No. 6117978  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: ROMMENS, JOHANNA M  
APPLICANT: FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TO ALZHEIMER'S DISEASE  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
STREET: High Street Tower - 125 High Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/124,698  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/592,541  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-124-698-4  
Query Match 20.8%; Score 97; DB 3; Length 467;  
Best Local Similarity 100.0%; Pred. No. 3.2e-84;  
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 213 IHMKGPLRLOQAYLIMISALMALVFIRKLPMTAVLLIAVISYDLVAVLCPPKGPLRLV 272  
Db 213 IHMKGPLRLOQAYLIMISALMALVFIRKLPMTAVLLIAVISYDLVAVLCPPKGPLRLV 272  
QY 273 ETAOERNETLFPALISSTWVLVNMAGDEPAORRV 309  
Db 273 ETAOERNETLFPALISSTWVLVNMAGDEPAORRV 309  
RESULT 33  
US-09-127-480-4  
Sequence 4, Application US/09127480  
Patent No. 6194153  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: ROMMENS, JOHANNA M  
APPLICANT: FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TO ALZHEIMER'S DISEASE  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
STREET: High Street Tower - 125 High Street

CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/127,480  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,541  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-127-480-4  
Query Match 20.8%; Score 97; DB 4; Length 467;  
Best Local Similarity 100.0%; Pred. No. 3.2e-84;  
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 213 IHMKGPLRLOQAYLIMISALMALVFIRKLPMTAVLLIAVISYDLVAVLCPPKGPLRLV 272  
Db 213 IHMKGPLRLOQAYLIMISALMALVFIRKLPMTAVLLIAVISYDLVAVLCPPKGPLRLV 272  
QY 273 ETAOERNETLFPALISSTWVLVNMAGDEPAORRV 309  
Db 273 ETAOERNETLFPALISSTWVLVNMAGDEPAORRV 309  
RESULT 34  
US-08-496-841C-4  
Sequence 4, Application US/08496841C  
Patent No. 6210919  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: ROMMENS, JOHANNA M  
APPLICANT: FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TO ALZHEIMER'S DISEASE  
NUMBER OF SEQUENCES: 175  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby & Darby, PC  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/496,841C  
FILING DATE: 28-Jun-1995  
CLASSIFICATION: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Paul F. Fehlnier, Ph.D.  
REGISTRATION NUMBER: 35,135

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 527-7700  
TELEFAX: (212) 753-6237  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-08-496-841C-4

Query Match 20.8%; Score 97; DB 4; Length 467;  
Best Local Similarity 100.0%; Pred. No. 3.2e-84;  
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 213 IHMKGPLRQQAIVLIMISALMALVFIKYLEPTAWLLAIVISYDVLAVLCPKGPLRLMV 272  
213 IHMKGPLRQQAIVLIMISALMALVFIKYLEPTAWLLAIVISYDVLAVLCPKGPLRLMV 272

Qy 273 ETAOERNETLFPALISSTWVLVNMAGDPEAORRV 309  
273 ETAOERNETLFPALISSTWVLVNMAGDPEAORRV 309

Db 273 ETAOERNETLFPALISSTWVLVNMAGDPEAORRV 309

RESULT 35  
US-08-496-841C-136  
Sequence 136, Application US/08496841C  
Patent No. 6210919  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
ROMMENS, JOHANNA M  
FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TO ALZHEIMER'S DISEASE  
NUMBER OF SEQUENCES: 175  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby & Darby, PC  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/496,841C  
FILING DATE: 28-Jun-1995  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Paul F. Fehlner, Ph.D.  
REGISTRATION NUMBER: 35,135  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 527-7700  
TELEFAX: (212) 753-6237  
INFORMATION FOR SEQ ID NO: 136:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 136:  
US-08-496-841C-136

Query Match 20.8%; Score 97; DB 4; Length 467;  
Best Local Similarity 100.0%; Pred. No. 3.2e-84;  
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 213 IHMKGPLRQQAIVLIMISALMALVFIKYLEPTAWLLAIVISYDVLAVLCPKGPLRLMV 272

Db 213 IHMKGPLRQQAIVLIMISALMALVFIKYLEPTAWLLAIVISYDVLAVLCPKGPLRLMV 272  
213 IHMKGPLRQQAIVLIMISALMALVFIKYLEPTAWLLAIVISYDVLAVLCPKGPLRLMV 272

Qy 273 ETAOERNETLFPALISSTWVLVNMAGDPEAORRV 309  
273 ETAOERNETLFPALISSTWVLVNMAGDPEAORRV 309

Db 273 ETAOERNETLFPALISSTWVLVNMAGDPEAORRV 309

RESULT 36  
US-09-124-523-4  
Sequence 4, Application US/09124523  
Patent No. 6395960  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
ROMMENS, JOHANNA M  
FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TO ALZHEIMER'S DISEASE  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
STREET: High Street Tower - 125 High Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/124,523  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,541  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-124-523-4

Query Match 20.8%; Score 97; DB 4; Length 467;  
Best Local Similarity 100.0%; Pred. No. 3.2e-84;  
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 213 IHMKGPLRQQAIVLIMISALMALVFIKYLEPTAWLLAIVISYDVLAVLCPKGPLRLMV 272  
213 IHMKGPLRQQAIVLIMISALMALVFIKYLEPTAWLLAIVISYDVLAVLCPKGPLRLMV 272

Db 213 IHMKGPLRQQAIVLIMISALMALVFIKYLEPTAWLLAIVISYDVLAVLCPKGPLRLMV 272

Qy 273 ETAOERNETLFPALISSTWVLVNMAGDPEAORRV 309  
273 ETAOERNETLFPALISSTWVLVNMAGDPEAORRV 309

Db 273 ETAOERNETLFPALISSTWVLVNMAGDPEAORRV 309

RESULT 37  
US-08-788-231A-17  
Sequence 17, Application US/08788231A  
Patent No. 6019974  
GENERAL INFORMATION:  
APPLICANT: L'Hernault, Steven W.  
TITLE OF INVENTION: SPE-4 RELATED PEPTIDES, ANTIBODIES AND  
METHODS

NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/788,231A  
FILING DATE: 24-JAN-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/010,672  
FILING DATE: 26-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Feider, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 60-95  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 465 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: not relevant  
MOLECULE TYPE: Protein  
HYPOTHETICAL: YES  
US-08-788-231A-17

Query Match 18.0%; Score 84; DB 3; Length 465;  
Best Local Similarity 100.0%; Pred. No. 7.9e-72;  
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 QVQDDEDEDELTKYAKVIMLFVPTLCMVVYVATISVSFTYRKQGLIYTPTE 120  
DB 60 QVQDDEDEDELTKYAKVIMLFVPTLCMVVYVATISVSFTYRKQGLIYTPTE 119  
QY 121 DTEVQGRALHSILNAIMISIV 144  
DB 120 DTEVQGRALHSILNAIMISIV 143

RESULT 38  
US-08-967-101-138  
Sequence 138, Application US/08967101  
Patent No. 5840540  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: ROMMENS, JOHANNA M  
APPLICANT: FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
STREET: High Street Tower - 125 High Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/967,101  
FILING DATE: 10-NOV-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/592,541  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7100  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 138:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 372 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-967-101-138

Query Match 6.0%; Score 28; DB 2; Length 372;  
Best Local Similarity 100.0%; Pred. No. 1.5e-18;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 ISVYDLVAVLCPKGRLMIVETAQRNE 280  
DB 183 ISVYDLVAVLCPKGRLMIVETAQRNE 210

RESULT 39  
US-08-592-541-138  
Sequence 138, Application US/08592541  
Patent No. 5986054  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: ROMMENS, JOHANNA M  
APPLICANT: FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
STREET: High Street Tower - 125 High Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,541  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 138:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 372 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-592-541-138

Query Match 6.0%; Score 28; DB 2; Length 372;  
Best Local Similarity 100.0%; Pred. No. 1.5e-18;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 253 ISVYDLVAVLCRPGPLRMVETAEERNE 280  
|||||  
Db 183 ISVYDLVAVLCRPGPLRMVETAEERNE 210

## RESULT 40

US-09-124-698-138

Sequence 138, Application US/09124698

Patent No. 6117978

GENERAL INFORMATION:

APPLICANT: ST. GEORGE-HYSLOP, PETER H

APPLICANT: ROMMENS, JOHANNA M

APPLICANT: FRASER, PAUL E

TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED

NUMBER OF SEQUENCES: 183

CORRESPONDENCE ADDRESS:

ADDRESSEE: TESTA, HURWITZ &amp; THIBEAULT

STREET: High Street Tower - 125 High Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/124,698

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/592,541

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Pitcher, Edmund R.

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 138:

SEQUENCE CHARACTERISTICS:

LENGTH: 372 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-124-698-138

Query Match 6.0%; Score 28; DB 3; Length 372;  
Best Local Similarity 100.0%; Pred. No. 1.5e-18;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 253 ISVYDLVAVLCRPGPLRMVETAEERNE 280  
|||||  
Db 183 ISVYDLVAVLCRPGPLRMVETAEERNE 210

## RESULT 41

US-09-127-480-138

Sequence 138, Application US/09127480

Patent No. 6194153

GENERAL INFORMATION:

APPLICANT: ST. GEORGE-HYSLOP, PETER H

APPLICANT: ROMMENS, JOHANNA M

APPLICANT: FRASER, PAUL E

TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED

NUMBER OF SEQUENCES: 183

CORRESPONDENCE ADDRESS:

ADDRESSEE: TESTA, HURWITZ &amp; THIBEAULT

STREET: High Street Tower - 125 High Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/496,841C

FILING DATE: 28-Jun-1995

CLASSIFICATION: &lt;unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: Paul F. Fehner, Ph.D.

REGISTRATION NUMBER: 35,135

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 527-7700

TELEFAX: (212) 753-6237

STREET: High Street Tower - 125 High Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/127,480

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/592,541

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Pitcher, Edmund R.

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 138:

SEQUENCE CHARACTERISTICS:

LENGTH: 372 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-127-480-138

Query Match 6.0%; Score 28; DB 4; Length 372;  
Best Local Similarity 100.0%; Pred. No. 1.5e-18;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 253 ISVYDLVAVLCRPGPLRMVETAEERNE 280  
|||||  
Db 183 ISVYDLVAVLCRPGPLRMVETAEERNE 210

## RESULT 42

US-08-496-841C-138

Sequence 138, Application US/08496841C

Patent No. 6210919

GENERAL INFORMATION:

APPLICANT: ST. GEORGE-HYSLOP, PETER H

APPLICANT: ROMMENS, JOHANNA M

APPLICANT: FRASER, PAUL E

TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED

NUMBER OF SEQUENCES: 175

CORRESPONDENCE ADDRESS:

ADDRESSEE: Darby &amp; Darby, PC

STREET: 805 Third Avenue

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/496,841C

FILING DATE: 28-Jun-1995

CLASSIFICATION: &lt;unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: Paul F. Fehner, Ph.D.

REGISTRATION NUMBER: 35,135

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 527-7700

TELEFAX: (212) 753-6237

Thu Feb 6 15:36:10 2003

INFORMATION FOR SEQ ID NO: 138:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 372 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 138:  
US-08-496-841C-138

Query Match 6.0%; Score 28; DB 4; Length 372;  
Best Local Similarity 100.0%; Pred. No. 1.5e-18;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 253 ISYDVLAVLCPKGPLRLMVLVETAEERNE 280  
|||||  
Db 183 ISYDVLAVLCPKGPLRLMVLVETAEERNE 210

RESULT 43

US-09-124-523-138  
Sequence 138, Application US/09124523  
Patent No. 6395960  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: ROMMENS, JOHANNA M  
APPLICANT: FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
STREET: High Street Tower - 125 High Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/124,523  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,541  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 138:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 372 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-124-523-138

Query Match 6.0%; Score 28; DB 4; Length 372;  
Best Local Similarity 100.0%; Pred. No. 1.5e-18;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 253 ISYDVLAVLCPKGPLRLMVLVETAEERNE 280  
|||||  
Db 183 ISYDVLAVLCPKGPLRLMVLVETAEERNE 210

RESULT 44

US-08-875-972-29  
Sequence 29, Application US/08875972  
Patent No. 598564  
GENERAL INFORMATION:  
APPLICANT: Huntington Potter and Jinhue Li  
TITLE OF INVENTION: ASSAY FOR IDENTIFYING GENES CAUSING  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.  
STREET: Two Millita Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173-4799

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/875,972  
FILING DATE: 08-AUG-97  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/002,448  
FILING DATE: 16-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan Esq., Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: H095-03PA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781) 861-6240  
TELEFAX: (781) 861-9540  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 447 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-875-972-29

Query Match 6.0%; Score 28; DB 2; Length 447;  
Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 253 ISYDVLAVLCPKGPLRLMVLVETAEERNE 280  
|||||  
Db 259 ISYDVLAVLCPKGPLRLMVLVETAEERNE 286

RESULT 45

US-08-967-101-137  
Sequence 137, Application US/08967101  
Patent No. 5840540  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: ROMMENS, JOHANNA M  
APPLICANT: FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
STREET: High Street Tower - 125 High Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/967,101  
FILING DATE: 10-NOV-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/592,541  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 137:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 448 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-967-101-137

Query Match 6.0%; Score 28; DB 2; Length 448;  
Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 ISVYDLVAVLCPKGPLRMLVETAQERNE 280  
Db 259 ISVYDLVAVLCPKGPLRMLVETAQERNE 286

RESULT 46  
US-08-592-541-137  
Sequence 137, Application US/08592541  
Patent No. 5986054  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: ROMMENS, JOHANNA M  
APPLICANT: FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TO ALZHEIMER'S DISEASE  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
STREET: High Street Tower - 125 High Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,541  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 137:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 448 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-592-541-137

Query Match 6.0%; Score 28; DB 2; Length 448;  
Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 ISVYDLVAVLCPKGPLRMLVETAQERNE 280  
Db 259 ISVYDLVAVLCPKGPLRMLVETAQERNE 286

RESULT 47  
US-08-888-077A-19  
Sequence 19, Application US/08888077A  
Patent No. 6020143  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: ROMMENS, JOHANNA M  
APPLICANT: FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TO ALZHEIMER'S DISEASE AND USES THEREFOR.  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLZ & MENTLIK  
STREET: 600 SOUTH AVENUE WEST  
CITY: WESTFIELD  
STATE: NJ  
COUNTRY: USA  
ZIP: 07090-1497  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/888,077A  
FILING DATE: 03-JUL-1997  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/592,541  
FILING DATE: 26-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: PALISI, THOMAS M  
REGISTRATION NUMBER: 36,629  
REFERENCE/DOCKET NUMBER: SCHERING 3.0-017 CIP CIP IV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 654-5000  
TELEFAX: (908) 654-7866  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 448 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-888-077A-19

Query Match 6.0%; Score 28; DB 3; Length 448;  
Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 ISVYDLVAVLCPKGPLRMLVETAQERNE 280  
Db 259 ISVYDLVAVLCPKGPLRMLVETAQERNE 286

RESULT 48  
US-09-124-698-137  
Sequence 137, Application US/09124698  
Patent No. 6117978  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: ROMMENS, JOHANNA M  
APPLICANT: FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TO ALZHEIMER'S DISEASE

NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
STREET: High Street Tower - 125 High Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/124,698  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/592,541  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 137:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 448 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-124-698-137

Query Match 6.0%; Score 28; DB 3; Length 448;  
Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 ISVYDLVAVLCPKGPLRLMVLVETAEGRNE 280  
|||||  
Db 259 ISVYDLVAVLCPKGPLRLMVLVETAEGRNE 286

RESULT 49  
US-09-127-480-137  
Sequence 137, Application US/09127480  
Patent No. 6194153  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: KOMMENS, JOHANNA M  
APPLICANT: FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
STREET: High Street Tower - 125 High Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/127,480  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,541  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 137:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 448 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-127-480-137

Query Match 6.0%; Score 28; DB 4; Length 448;  
Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 ISVYDLVAVLCPKGPLRLMVLVETAEGRNE 280  
|||||  
Db 259 ISVYDLVAVLCPKGPLRLMVLVETAEGRNE 286

RESULT 50  
US-09-227-725A-2  
Sequence 2, Application US/09227725A  
Patent No. 6383758  
GENERAL INFORMATION:  
APPLICANT: St. George-Hyslop, Peter H.  
APPLICANT: Kommens, Johanna  
APPLICANT: Fraser, Paul E.  
TITLE OF INVENTION: Alzheimer's Related Proteins and Methods  
TITLE OF INVENTION: of use  
FILE REFERENCE: 1034/1P810-US1  
CURRENT APPLICATION NUMBER: US/09/227,725A  
CURRENT FILING DATE: 1999-01-08  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 448  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-09-227-725A-2

Query Match 6.0%; Score 28; DB 4; Length 448;  
Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 ISVYDLVAVLCPKGPLRLMVLVETAEGRNE 280  
|||||  
Db 259 ISVYDLVAVLCPKGPLRLMVLVETAEGRNE 286

Search completed: February 6, 2003, 14:21:24  
Job time : 41.1598 secs





84	28	6.0	359	22	AAW23821	Human EST encoded
85	28	6.0	447	17	AAW28508	Full AD/AD3LP seq
86	28	6.0	448	17	AAW05762	Human presentin-2
87	28	6.0	448	17	AAW05763	Presentin-2 M33V
88	28	6.0	448	17	AAW05764	Presentin-2 M41I
89	28	6.0	448	17	AAW05765	Presentin-2 I420T
90	28	6.0	448	18	AAW11321	Human AD4 protein.
91	28	6.0	448	19	AAW23967	Human presentin-2
92	28	6.0	448	20	AAW23898	Human presentin-2
93	28	6.0	448	21	AAW07972	Amino acid sequenc
94	28	6.0	448	22	AAE10799	Human presentin-2
95	28	6.0	448	22	AAE10799	Human presentin-2
96	28	6.0	448	22	AAE10799	Amino acid sequenc
97	28	6.0	448	22	AAE10799	Human presentin-2
98	28	6.0	448	23	AAW09417	Human presentin-2
99	28	6.0	448	23	AAE17049	Human mutant prese
100	28	6.0	448	23	AAE17052	Human mutant prese
101	28	6.0	448	23	AAE17052	Human mutant prese
102	28	6.0	448	23	AAE17052	Human mutant prese
103	28	6.0	448	23	AAE17052	Human mutant prese
104	28	6.0	448	23	AAE17052	Human mutant prese
105	28	6.0	448	23	AAE17052	Human mutant prese
106	28	6.0	448	23	AAE17052	Human mutant prese
107	28	6.0	448	23	AAE17052	Human mutant prese
108	28	6.0	448	23	AAE17052	Human mutant prese
109	28	6.0	448	23	AAE17052	Human mutant prese
110	28	6.0	448	23	AAE17052	Human mutant prese
111	28	6.0	448	23	AAE17052	Human mutant prese
112	28	6.0	448	23	AAE17052	Human mutant prese
113	28	6.0	448	23	AAE17052	Human mutant prese
114	28	6.0	448	23	AAE17052	Human mutant prese
115	28	6.0	448	23	AAE17052	Human mutant prese
116	28	6.0	448	23	AAE17052	Human mutant prese
117	28	6.0	448	23	AAE17052	Human mutant prese
118	28	6.0	448	23	AAE17052	Human mutant prese
119	28	6.0	448	23	AAE17052	Human mutant prese
120	28	6.0	448	23	AAE17052	Human mutant prese
121	28	6.0	448	23	AAE17052	Human mutant prese
122	28	6.0	448	23	AAE17052	Human mutant prese
123	28	6.0	448	23	AAE17052	Human mutant prese
124	28	6.0	448	23	AAE17052	Human mutant prese
125	28	6.0	448	23	AAE17052	Human mutant prese
126	28	6.0	448	23	AAE17052	Human mutant prese
127	28	6.0	448	23	AAE17052	Human mutant prese
128	28	6.0	448	23	AAE17052	Human mutant prese
129	28	6.0	448	23	AAE17052	Human mutant prese
130	28	6.0	448	23	AAE17052	Human mutant prese
131	28	6.0	448	23	AAE17052	Human mutant prese
132	28	6.0	448	23	AAE17052	Human mutant prese
133	28	6.0	448	23	AAE17052	Human mutant prese
134	28	6.0	448	23	AAE17052	Human mutant prese
135	28	6.0	448	23	AAE17052	Human mutant prese
136	28	6.0	448	23	AAE17052	Human mutant prese
137	28	6.0	448	23	AAE17052	Human mutant prese
138	28	6.0	448	23	AAE17052	Human mutant prese
139	28	6.0	448	23	AAE17052	Human mutant prese
140	28	6.0	448	23	AAE17052	Human mutant prese
141	28	6.0	448	23	AAE17052	Human mutant prese
142	28	6.0	448	23	AAE17052	Human mutant prese
143	28	6.0	448	23	AAE17052	Human mutant prese
144	28	6.0	448	23	AAE17052	Human mutant prese
145	28	6.0	448	23	AAE17052	Human mutant prese
146	28	6.0	448	23	AAE17052	Human mutant prese
147	28	6.0	448	23	AAE17052	Human mutant prese
148	28	6.0	448	23	AAE17052	Human mutant prese
149	28	6.0	448	23	AAE17052	Human mutant prese
150	28	6.0	448	23	AAE17052	Human mutant prese
151	28	6.0	448	23	AAE17052	Human mutant prese
152	28	6.0	448	23	AAE17052	Human mutant prese
153	28	6.0	448	23	AAE17052	Human mutant prese
154	28	6.0	448	23	AAE17052	Human mutant prese
155	28	6.0	448	23	AAE17052	Human mutant prese
156	28	6.0	448	23	AAE17052	Human mutant prese

230	7	1.5	25	22	ABR22871	Protein #4870 enco	303	7	1.5	159	18	AAW11786	Early onset Alzhei
231	7	1.5	48	20	AAV11731	Human 5' EST secre	304	7	1.5	166	22	ABG27891	Novel human diagno
232	7	1.5	48	20	AAW24168	Human 5' EST encoded	305	7	1.5	167	20	AAV36759	Amino acid sequenc
233	7	1.5	50	21	ACG28663	Arabidopsis thalia	306	7	1.5	168	21	AAV25333	Arabidopsis thalia
234	7	1.5	50	21	AAV59419	Arabidopsis thalia	307	7	1.5	172	14	AAV3569	Arabidopsis thalia
235	7	1.5	54	23	ABP03101	Human ORFX protein	308	7	1.5	173	14	AAV42126	Arabidopsis 18KD o
236	7	1.5	55	22	ABR38483	Peptide #5989 enco	309	7	1.5	173	19	AAV98470	A. thaliana oleosl
237	7	1.5	55	22	ABR23639	Protein #5638 enco	310	7	1.5	173	19	AAV64206	A. thaliana oleosl
238	7	1.5	55	22	AAV59098	Human brain expres	311	7	1.5	178	22	ABR65236	Drosophila melanog
239	7	1.5	55	22	AAV71633	Human bone marrow	312	7	1.5	181	22	AAV62374	S. epidermidis ope
240	7	1.5	55	22	AAV19243	Peptide #5677 enco	313	7	1.5	186	19	AAV57892	Protein of Clone A
241	7	1.5	55	22	AAV31927	Peptide #5964 enco	314	7	1.5	187	17	AAV38471	B. napus oleosin.
242	7	1.5	55	22	ABG41447	Human peptide enco	315	7	1.5	187	20	AAV33828	Amino acid sequenc
243	7	1.5	57	22	AAO04025	Human peptide enco	316	7	1.5	188	21	AAV30849	Arabidopsis thalia
244	7	1.5	58	19	AAV51842	Human polypeptide	317	7	1.5	188	21	AAV51131	Arabidopsis thalia
245	7	1.5	58	22	AAU22441	Rana temporaria te	318	7	1.5	202	21	AAV16444	Arabidopsis thalia
246	7	1.5	61	19	AAV51841	Human cardiovascu	319	7	1.5	213	23	ABP38711	Pinus radiata pero
247	7	1.5	61	19	AAV51843	Rana temporaria te	320	7	1.5	218	23	ABR47943	Staphylococcus epi
248	7	1.5	61	21	AAV18737	A temporin G precu	321	7	1.5	232	23	ABP60683	Listeria monocytog
249	7	1.5	63	22	AAV75233	Human colon cancer	322	7	1.5	234	17	AAV93802	Phaseolin promoter
250	7	1.5	64	19	AAV83938	Human secreted pro	323	7	1.5	238	22	ABR65748	B. napus oleosin.
251	7	1.5	67	23	ABG44153	Human peptide enco	324	7	1.5	241	22	AAU04044	Drosophila melanog
252	7	1.5	69	20	AAV11732	Human 5' EST secre	325	7	1.5	242	22	ABR30277	Streptococcus coel
253	7	1.5	70	22	ABR28992	Peptide #1643 enco	326	7	1.5	242	22	ABR35443	Peptide #2928 enco
254	7	1.5	70	22	ABR34160	Peptide #1666 enco	327	7	1.5	242	22	ABR20882	Peptide #2949 enco
255	7	1.5	70	22	ABR19604	Protein #1603 enco	328	7	1.5	242	22	AAV56286	Human brain expres
256	7	1.5	70	22	AAV54953	Human brain expres	329	7	1.5	242	22	AAV56286	Human brain expres
257	7	1.5	70	22	AAV67334	Human bone marrow	330	7	1.5	242	22	AAV68643	Human bone marrow
258	7	1.5	70	22	AAV51568	Peptide #1602 enco	331	7	1.5	242	22	AAV16455	Peptide #2889 enco
259	7	1.5	70	22	AAV27625	Peptide #1662 enco	332	7	1.5	242	22	AAV28950	Peptide #2987 enco
260	7	1.5	70	22	AAV02906	Peptide #1588 enco	333	7	1.5	242	22	AAV04184	Peptide #2866 enco
261	7	1.5	70	23	ABG36973	Human peptide enco	334	7	1.5	247	23	ABG38232	Human peptide enco
262	7	1.5	72	22	AAO01477	Human polypeptide	335	7	1.5	254	19	AAV32684	C glutamicum prote
263	7	1.5	82	23	ABP00848	Human ORFX protein	336	7	1.5	263	22	AAV64209	Oleoin-metallothi
264	7	1.5	85	22	AAV58947	Protonibacterium	337	7	1.5	264	22	AAV87751	Human T2R26 amino
265	7	1.5	87	20	AAV37867	Amino acid sequenc	338	7	1.5	274	21	AAV81752	Human AFP protein
266	7	1.5	88	22	ABG27884	Novel human diagno	339	7	1.5	276	22	AAV946974	Arabidopsis thalia
267	7	1.5	90	22	ABG29361	Novel human diagno	340	7	1.5	283	23	ABR05673	Human protein sequ
268	7	1.5	91	22	AAV89181	Human immune/haema	341	7	1.5	284	21	AAV39054	Human amygdala der
269	7	1.5	93	21	AAV44010	Zea mays protein f	342	7	1.5	286	22	AAV30808	Arabidopsis thalia
270	7	1.5	94	20	AAV48491	Human breast tumou	343	7	1.5	292	21	AAV81703	Amino acid sequenc
271	7	1.5	99	22	ABG29226	Novel human diagno	344	7	1.5	292	21	AAV81752	Streptococcus pneu
272	7	1.5	101	22	AAV66917	Human immune/haema	345	7	1.5	295	21	AAV30848	Arabidopsis thalia
273	7	1.5	107	22	AAV41658	Protonibacterium	346	7	1.5	295	21	AAV51130	Arabidopsis thalia
274	7	1.5	110	21	AAV32963	Pinus radiata tran	347	7	1.5	297	21	AAV06973	Arabidopsis thalia
275	7	1.5	114	23	AAU12983	Human hexokinase p	348	7	1.5	299	22	AAV87751	Human T2R21 amino
276	7	1.5	117	19	AAV64211	Oleoin-protein A	349	7	1.5	299	22	AAV81748	Herbically activ
277	7	1.5	117	22	AAV01922	Human polypeptide	350	7	1.5	300	22	ABV09612	Novel human diagno
278	7	1.5	118	19	AAV64207	Oleoin-hirudin fu	351	7	1.5	301	22	AAV1276	C glutamicum prote
279	7	1.5	118	20	AAV33829	Amino acid sequenc	352	7	1.5	302	22	ABV08067	Novel human diagno
280	7	1.5	118	23	ABP60681	Phaseolin promoter	353	7	1.5	307	21	AAV21066	Arabidopsis thalia
281	7	1.5	118	23	ABP60686	Phaseolin promoter	354	7	1.5	310	22	AAV76840	Corynebacterium gl
282	7	1.5	118	23	ABP60690	Phaseolin promoter	355	7	1.5	319	22	AAV21968	C glutamicum prote
283	7	1.5	118	23	ABP60692	Promoter-Oleosin t	356	7	1.5	327	22	AAV72266	Human olfactory re
284	7	1.5	118	23	ABP01341	Human ORFX protein	357	7	1.5	330	21	AAV39053	Arabidopsis thalia
285	7	1.5	119	21	AAV34728	Human secreted pro	358	7	1.5	333	21	AAV31065	Arabidopsis thalia
286	7	1.5	119	21	AAV82453	Human TGC-440 secr	359	7	1.5	334	21	AAV14646	Arabidopsis thalia
287	7	1.5	119	21	AAV73737	Human signal pepti	360	7	1.5	339	22	ABV16430	Novel human diagno
288	7	1.5	119	21	AAV66668	Membrane-bound pro	361	7	1.5	339	23	AAV49044	Novel human diagno
289	7	1.5	119	22	AAV29093	Human PRO polypept	362	7	1.5	341	22	AAV1074	C glutamicum prote
290	7	1.5	119	22	AAV63977	Amino acid sequenc	363	7	1.5	346	22	ABV1976	Novel human diagno
291	7	1.5	119	22	AAV87538	Human PRO842. Hom	364	7	1.5	361	22	ABV58074	Drosophila melanog
292	7	1.5	119	22	AAV65191	Human PRO842 (UNQ	365	7	1.5	361	22	ABV57345	Drosophila melanog
293	7	1.5	121	22	ABG26796	Novel human diagno	366	7	1.5	364	22	ABV62333	Mouse ischaemic co
294	7	1.5	126	22	AAV01854	Human polypeptide	367	7	1.5	366	22	AAV66863	Drosophila melanog
295	7	1.5	130	22	AAV88369	Human immune/haema	368	7	1.5	368	21	AAV21064	Arabidopsis oleosl
296	7	1.5	139	22	ABV68552	Drosophila melanog	369	7	1.5	368	21	AAV34286	Arabidopsis thalia
297	7	1.5	141	19	AAV58386	Human secreted pro	370	7	1.5	376	23	ABV49919	Listeria monocytog
298	7	1.5	141	22	AAV90685	Human BR595_4 prot	371	7	1.5	392	19	AAV48391	Homo sapiens pancr
299	7	1.5	144	21	AAV09686	Arabidopsis thalia	372	7	1.5	392	20	AAV04120	Pancreas derived p
300	7	1.5	148	22	AAV01988	Human polypeptide	373	7	1.5	392	23	AAV14266	Human Pancreas-der
301	7	1.5	154	17	AAV88472	A. thaliana oleosl	374	7	1.5	394	8	AAV70434	Amino acids 71-464
302	7	1.5	155	22	ABG21215	Novel human diagno	375	7	1.5	394	11	AAV04834	N-myc protein. A

376	7	1.5	401	23	ABR93450	Herbicidally activ	449	7	1.5	712	18	AAW30749	Rat YTS21 gene pro
377	7	1.5	403	21	AA637514	Arabidopsis thalia	450	7	1.5	714	22	ABG12681	Novel human diagno
378	7	1.5	405	20	AA637514	Pancreas derived p	451	7	1.5	731	22	ABG12985	Novel human diagno
379	7	1.5	405	20	AA637514	Protein encoded by	452	7	1.5	731	23	ABG92517	Herbicidally activ
380	7	1.5	405	21	AA637513	Arabidopsis thalia	453	7	1.5	739	22	AA693019	C glutamicum prote
381	7	1.5	405	23	AA614267	Human Pancreas-der	454	7	1.5	740	19	AAW61533	Human Fas-binding
382	7	1.5	410	21	AA651721	Human secreted pro	455	7	1.5	740	22	AAU00631	Human death-assoc
383	7	1.5	410	21	AA651722	Human secreted pro	456	7	1.5	740	22	AAU00631	Human death-assoc
384	7	1.5	411	21	AA639052	Arabidopsis thalia	457	7	1.5	746	21	AA682742	DNA replication an
385	7	1.5	415	22	AB652991	Escherichia coli p	458	7	1.5	748	22	AB657847	Drosophila melanog
386	7	1.5	418	22	AB610859	Mycobacterium tube	459	7	1.5	748	22	AAU34685	E. coli cellular p
387	7	1.5	418	23	ABP61052	Lactobacillus rham	460	7	1.5	748	22	AAU34685	Salmonella typhi c
388	7	1.5	425	22	AB671352	Arabidopsis thalia	461	7	1.5	765	22	AAU03520	Human protein kina
389	7	1.5	434	21	AA672840	Human adipophilin-	462	7	1.5	765	23	AAE16275	Human kinase PKIN-
390	7	1.5	438	22	AB62864	Escherichia coli p	463	7	1.5	774	19	AAW80411	A secreted protein
391	7	1.5	443	21	AA614645	Arabidopsis thalia	464	7	1.5	774	23	ABP61829	Human polypeptide
392	7	1.5	443	22	AA620005	Arabidopsis 3-keto	465	7	1.5	775	23	AAU93161	Human polypeptide
393	7	1.5	446	21	AA614644	Arabidopsis thalia	466	7	1.5	794	22	AAU39245	Arabidopsis trans
394	7	1.5	447	21	AA637512	Arabidopsis thalia	467	7	1.5	808	22	AAW41031	Human polypeptide
395	7	1.5	449	22	AB650668	C. elegans UNC-5 p	468	7	1.5	814	22	ABG21977	Human polypeptide
396	7	1.5	451	23	ABP60688	Phaseolin promoter	469	7	1.5	817	22	ABG95119	Novel human diagno
397	7	1.5	452	21	AA605734	Arabidopsis thalia	470	7	1.5	830	23	ABR04808	Human protein sequ
398	7	1.5	456	21	AA605733	Arabidopsis thalia	471	7	1.5	830	23	ABR04810	LDL receptor bindi
399	7	1.5	461	18	AAW19276	Pseudomonas aerugi	472	7	1.5	830	23	ABR04811	LDL receptor bindi
400	7	1.5	477	18	AAW25153	Nsp7524111 restric	473	7	1.5	830	23	ABR04812	LDL receptor bindi
401	7	1.5	478	17	AA695956	Eukaryotic cell gr	474	7	1.5	830	23	ABR04813	LDL receptor bindi
402	7	1.5	482	21	AA658366	Anabaena variabil	475	7	1.5	849	23	ABR97265	Novel human protei
403	7	1.5	489	19	ABG64061	Human albumin fusi	476	7	1.5	896	19	AAW41252	Xenopus paraxial p
404	7	1.5	494	19	AAW42394	Pyrococcus furiosu	477	7	1.5	903	20	AAW28713	Deactivated yeast m
405	7	1.5	498	12	AA613713	PRV glycoprotein g	478	7	1.5	906	20	AAW28713	Yeast multifunctio
406	7	1.5	498	12	AA613916	Pseudorabies virus	479	7	1.5	906	20	AAW28712	Mutant yeast multi
407	7	1.5	502	22	ABR78823	Pseudorabies virus	480	7	1.5	906	22	AAW20180	Candida tropicalis
408	7	1.5	505	22	AAE01380	Human gene 4 encod	481	7	1.5	906	22	AAW20181	C. tropicalis multi
409	7	1.5	507	22	AB660080	Human breast cance	482	7	1.5	906	22	AAW20182	C. tropicalis multi
410	7	1.5	512	23	ABP41659	Human ovarian anti	483	7	1.5	906	22	AAW20183	C. tropicalis multi
411	7	1.5	525	21	AA625284	Eucalyptus grandis	484	7	1.5	938	22	ABR60562	Drosophila melanog
412	7	1.5	527	20	AAW29515	Human lung tumour	485	7	1.5	978	22	ABR60562	Drosophila melanog
413	7	1.5	527	21	AA644492	Human lung tumour-	486	7	1.5	1022	22	ABR62978	Drosophila melanog
414	7	1.5	527	22	AAE13805	Human lung tumour-	487	7	1.5	1024	17	AAW03184	Rice Xa21 (RRK-F)
415	7	1.5	527	22	AAU29284	Human PRO polypept	488	7	1.5	1062	22	ABR71830	Drosophila melanog
416	7	1.5	527	23	AAU77927	Human drug-metabol	489	7	1.5	1134	22	AAU28072	Novel human secret
417	7	1.5	527	23	AAE15434	Human drug-metabol	490	7	1.5	1135	22	ABR62435	Drosophila melanog
418	7	1.5	530	23	AAW52210	Bacillus cellular dif	491	7	1.5	1149	22	AAW41048	Human polypeptide
419	7	1.5	542	17	AA690772	Oleosin steatocoe	492	7	1.5	1159	22	ABR63460	Drosophila melanog
420	7	1.5	545	20	AA63830	Drosophila melanog	493	7	1.5	1162	21	AAW96255	Kaposi's sarcoma-a
421	7	1.5	545	22	ABR68867	Drosophila melanog	494	7	1.5	1162	21	AAW96255	HIV8 ORF 73 protei
422	7	1.5	545	22	AAU38979	Drosophila G-prote	495	7	1.5	1162	22	AAW62331	Amino acid sequenc
423	7	1.5	559	22	AAW38813	Human polypeptide	496	7	1.5	1162	22	ABR62331	Kaposi's sarcoma-a
424	7	1.5	564	23	ABP40294	Staphylococcus epi	497	7	1.5	1174	22	AAW85039	Shrimp white spot
425	7	1.5	566	21	AAW66641	Membrane-bound pro	498	7	1.5	1263	22	AAW18244	Plasmodium falci
426	7	1.5	566	22	AAU29051	Human PRO polypept	499	7	1.5	1263	22	AAW94749	Human protein sequ
427	7	1.5	566	22	ABR65164	Human PRO710 (UNC3	500	7	1.5	1287	22	ABR44552	Human wound healin
428	7	1.5	571	23	ABP39020	Staphylococcus epi	501	7	1.5	1352	21	AAW31776	Arabidopsis thalia
429	7	1.5	580	22	ABR62193	Drosophila melanog	502	7	1.5	1371	22	ABG21216	Novel human diagno
430	7	1.5	585	22	ABG16868	Novel human diagno	503	7	1.5	1406	21	AAW31775	Arabidopsis thalia
431	7	1.5	596	21	AAW39850	Arabidopsis thalia	504	7	1.5	1407	21	AAW31774	Arabidopsis thalia
432	7	1.5	611	22	ABR60610	Drosophila melanog	505	7	1.5	1413	23	AAE21729	Human PKIN-24 prot
433	7	1.5	613	22	ABG22924	Novel human diagno	506	7	1.5	1488	22	AAE10801	Rat TGF-beta Resis
434	7	1.5	616	23	ABR83078	Synoviolin related	507	7	1.5	1489	22	AAE10803	Mouse TGF-beta Res
435	7	1.5	617	22	AAE01355	Human gene 4 encod	508	7	1.5	1490	22	AAE10802	Human TGF-beta Res
436	7	1.5	617	23	ABR83074	Synoviolin amino a	509	7	1.5	1572	18	AAW27160	Mouse receptor ME2
437	7	1.5	617	23	ABG64063	Human albumin fusi	510	7	1.5	1651	23	ABG66725	Human novel polype
438	7	1.5	624	22	ABR70398	Drosophila melanog	511	7	1.5	1675	21	AAW42658	Human OREF ORE2422
439	7	1.5	641	22	AAW23689	Human EST encoded	512	7	1.5	1905	22	ABR59243	Drosophila melanog
440	7	1.5	645	22	AAW40599	Human polypeptide	513	7	1.5	1905	22	AAU38925	Drosophila G-prote
441	7	1.5	659	23	ABR93561	Herbicidally activ	514	7	1.5	1972	19	AAW81171	Human BAZ2-Beta pr
442	7	1.5	664	22	AAW39262	Human polypeptide	515	7	1.5	2047	22	AAU75883	Human adhesion mol
443	7	1.5	671	23	ABR92516	Herbicidally activ	516	7	1.5	2175	22	ABR65698	Drosophila melanog
444	7	1.5	687	22	ABR59384	Arabidopsis seed s	517	7	1.5	2278	21	AAW53677	Sequence g1/341388
445	7	1.5	689	20	AAW42698	Arabidopsis seed s	518	7	1.5	2281	22	AAW40857	Human polypeptide
446	7	1.5	689	21	AAW01673	F151 protein sequ	519	7	1.5	2380	21	AAW18515	Plasmodium falci
447	7	1.5	689	21	AAW57036	Fertilisation-lnde	520	7	1.5	2415	22	ABG20279	Novel human diagno
448	7	1.5	689	22	AAW82455	Arabidopsis seed-s	521	7	1.5	2448	22	AAU36277	Pseudomonas aerugi



668	6	1.3	46	22	AAM24401	Human EST encoded	741	6	1.3	58	22	AAG73350	Human gene 5'-end
669	6	1.3	46	22	AAU14457	Human novel protei	742	6	1.3	58	22	AAU73386	Human gene 5'-end
670	6	1.3	48	19	AAW75157	Human secreted pro	743	6	1.3	58	23	ABG64205	Human albumin fusi
671	6	1.3	48	22	ABH70642	Drosophila melanog	744	6	1.3	58	23	ABG64206	Human albumin fusi
672	6	1.3	50	20	AAZ52739	Human secreted pro	745	6	1.3	58	23	ABG44629	Human peptide enco
673	6	1.3	50	22	AAU56573	Propionibacterium	746	6	1.3	59	21	AAU17388	Arabidopsis thalia
674	6	1.3	50	22	AAU62745	Propionibacterium	747	6	1.3	59	22	AAU41209	Propionibacterium
675	6	1.3	50	22	ABH40235	Peptide #7741 enco	748	6	1.3	59	22	AAH82515	Human immune/haema
676	6	1.3	50	22	ABB24665	Protein #664 enco	749	6	1.3	59	22	AAO10963	Human polypeptide
677	6	1.3	50	22	AAH61021	Human brain expres	750	6	1.3	60	20	AAH99073	Immunoglobulin Jk
678	6	1.3	50	22	AAH73721	Human bone marrow	751	6	1.3	60	20	AAH88527	Fusion of killer t
679	6	1.3	50	22	AAO73773	Human polyepitide	752	6	1.3	60	22	AAU47728	Propionibacterium
680	6	1.3	50	22	AAH33908	Peptide #7945 enco	753	6	1.3	61	22	ABG15345	Novel human diagno
681	6	1.3	50	23	ABG43599	Human peptide enco	754	6	1.3	61	22	ABG15347	Novel human diagno
682	6	1.3	51	15	ABP07947	Human ORFX protein	755	6	1.3	61	22	AAU14221	Human novel protei
683	6	1.3	51	22	AAH59858	Hepatitis C virus	756	6	1.3	62	16	AAH69526	Anti-HCV antibody
684	6	1.3	51	22	AAH14820	Human nervous syst	757	6	1.3	62	16	AAH69531	Anti-HCV antibody
685	6	1.3	51	23	ABP02835	Human ORFX protein	758	6	1.3	62	16	AAH69535	Anti-HCV antibody
686	6	1.3	52	20	AAH73369	Human prostate tum	759	6	1.3	62	22	AAU53228	Propionibacterium
687	6	1.3	52	19	AAH75080	Human secreted pro	760	6	1.3	62	22	AAU67544	Propionibacterium
688	6	1.3	53	21	AAH38376	Human secreted pro	761	6	1.3	62	23	ABP06544	Human ORFX protein
689	6	1.3	53	21	AAH16526	Bacteriophage 77.7	762	6	1.3	63	22	ABH69523	Drosophila melanog
690	6	1.3	53	22	ABH15233	Human nervous syst	763	6	1.3	63	22	AAU67214	Propionibacterium
691	6	1.3	53	22	ABH03171	Human musculoskele	764	6	1.3	63	22	AAH80801	Human haematologic
692	6	1.3	53	23	AAH95379	Human calcium tran	765	6	1.3	63	22	AAH94282	Human reproductive
693	6	1.3	54	20	AAH10870	Amino acid sequenc	766	6	1.3	63	22	AAO13771	Human polypeptide
694	6	1.3	54	21	AAH39049	Human secreted pro	767	6	1.3	64	21	AAH57638	Arabidopsis thalia
695	6	1.3	54	22	ABH30738	Peptide #3389 enco	768	6	1.3	64	22	AAU39438	Propionibacterium
696	6	1.3	54	22	ABH35914	Peptide #420 enco	769	6	1.3	64	22	AAU62362	Propionibacterium
697	6	1.3	54	22	ABH21323	Protein #3322 enco	770	6	1.3	64	22	ABG05340	Novel human diagno
698	6	1.3	54	22	AAH56718	Human brain expres	771	6	1.3	64	23	ABP42478	Human ovarian anti
699	6	1.3	54	22	AAH69094	Human bone marrow	772	6	1.3	64	23	ABP31422	Human ORFP95 prote
700	6	1.3	54	22	AAH16932	Peptide #3366 enco	773	6	1.3	65	15	AAH59857	Hepatitis C virus
701	6	1.3	54	22	AAH29417	Peptide #3454 enco	774	6	1.3	65	20	AAH12573	Human 5' EST secre
702	6	1.3	54	22	AAH04625	Peptide #307 enco	775	6	1.3	65	21	AAH15090	Zea mays protein f
703	6	1.3	54	22	AAH01315	Human gene 6 enco	776	6	1.3	65	21	AAH78232	Human signal pepti
704	6	1.3	55	23	ABH38693	Human peptide enco	777	6	1.3	65	22	AAH57411	Propionibacterium
705	6	1.3	55	16	AAH82635	Human mucosal lym	778	6	1.3	65	22	ABH30691	Peptide #3342 enco
706	6	1.3	55	22	ABH5038	Peptide #2544 enco	779	6	1.3	65	22	ABH35865	Peptide #3371 enco
707	6	1.3	55	22	ABH38298	Peptide #5804 enco	780	6	1.3	65	22	AAH18568	Human lung antigen
708	6	1.3	55	22	ABH23479	Protein #5478 enco	781	6	1.3	65	22	AAH69047	Human bone marrow
709	6	1.3	55	22	AAH58920	Human brain expres	782	6	1.3	65	22	AAH87734	Human immune/haema
710	6	1.3	55	22	AAH71442	Human bone marrow	783	6	1.3	65	22	AAH16884	Peptide #3405 enco
711	6	1.3	55	22	AAH89110	Human immune/haema	784	6	1.3	65	22	AAH29368	Peptide #3405 enco
712	6	1.3	55	22	AAH19095	Peptide #5529 enco	785	6	1.3	65	23	ABH38646	Human peptide enco
713	6	1.3	55	22	AAH28545	Peptide #2582 enco	786	6	1.3	65	23	ABP32298	Human ORF1271 prot
714	6	1.3	55	22	AAH31734	Peptide #5771 enco	787	6	1.3	66	20	AAH100341	Human secreted pro
715	6	1.3	55	23	ABH37755	Human peptide enco	788	6	1.3	66	21	AAH38286	Human secreted pro
716	6	1.3	55	23	ABH41248	Human peptide enco	789	6	1.3	66	21	AAH38290	Human secreted pro
717	6	1.3	55	23	ABP07038	Human ORFX protein	790	6	1.3	66	23	ABP01476	Human ORFX protein
718	6	1.3	56	21	AAH26102	Zea mays protein f	791	6	1.3	67	15	AAH59856	Hepatitis C virus
719	6	1.3	56	22	AAU40524	Novel human diagno	792	6	1.3	68	21	AAH19568	Arabidopsis thalia
720	6	1.3	56	22	ABH26499	Propionibacterium	793	6	1.3	68	21	AAH45680	Arabidopsis thalia
721	6	1.3	56	22	ABH17932	Human nervous syst	794	6	1.3	68	22	AAH61452	Propionibacterium
722	6	1.3	56	22	AAH004488	Human polypeptide	795	6	1.3	68	22	AAH28249	Novel human diagno
723	6	1.3	56	22	AAH77220	Human colon cancer	796	6	1.3	68	22	AAH56683	Human reproductive
724	6	1.3	56	23	ABP32030	Human ORF1003 prot	797	6	1.3	68	23	ABP02110	Human ORFX protein
725	6	1.3	57	21	AAH52114	Human secreted pro	798	6	1.3	69	23	AAH12169	Arabidopsis thalia
726	6	1.3	57	22	ABH41035	Peptide #8541 enco	799	6	1.3	69	21	AAH12927	Arabidopsis thalia
727	6	1.3	57	22	AAH61890	Human brain expres	800	6	1.3	69	21	AAH36767	Arabidopsis thalia
728	6	1.3	57	22	AAH74692	Human bone marrow	801	6	1.3	69	22	AAH50241	Propionibacterium
729	6	1.3	57	22	AAH34808	Peptide #845 enco	802	6	1.3	69	22	AAH56620	Propionibacterium
730	6	1.3	57	22	AAH06350	Human foetal prote	803	6	1.3	69	22	AAO10555	Human polypeptide
731	6	1.3	57	23	ABP03087	Human ORFX protein	804	6	1.3	69	22	AAU20116	Human DNA repair a
732	6	1.3	57	23	ABP30795	Streptococcus poly	805	6	1.3	69	23	ABP05110	Human ORFX protein
733	6	1.3	58	22	ABH19930	Novel human diagno	806	6	1.3	70	22	AAU32751	Human ORFX secret
734	6	1.3	58	22	ABH41187	Peptide #8693 enco	807	6	1.3	70	23	ABP42746	Human ovarian anti
735	6	1.3	58	22	ABH25208	Protein #7207 enco	808	6	1.3	71	16	AAH69524	Anti-HCV antibody
736	6	1.3	58	22	AAH62047	Human brain expres	809	6	1.3	71	16	AAH69527	Anti-HCV antibody
737	6	1.3	58	22	AAH74849	Human bone marrow	810	6	1.3	71	16	AAH69532	Anti-HCV antibody
738	6	1.3	58	22	AAH84421	Human immune/haema	811	6	1.3	71	16	AAH69536	Anti-HCV antibody
739	6	1.3	58	22	AAH20434	Peptide #6868 enco	812	6	1.3	71	21	AAH19968	Arabidopsis thalia
740	6	1.3	58	22	AAH34966	Peptide #9003 enco	813	6	1.3	71	21	AAH37982	Arabidopsis thalia

814	6	1.3	71	22	ABB44369	Peptide #11875 enc	887	6	1.3	80	22	AAU42555	Protonibacterium
815	6	1.3	71	22	ABR27226	Protein #9225 enco	888	6	1.3	80	22	ABG14485	Novel human diagno
816	6	1.3	71	22	AAW65436	Human brain expres	889	6	1.3	80	22	AAW62696	Human immune/haema
817	6	1.3	71	22	AAW78110	Human bone marrow	890	6	1.3	80	23	AAU87818	Yeast ELO2 partial
818	6	1.3	71	22	AAO11264	Human polypeptide	891	6	1.3	80	23	AAU87819	Yeast ELO2 protein
819	6	1.3	71	22	AAW12191	Peptide #8425 enco	892	6	1.3	81	12	AAAR10311	Ovary tissue trans
820	6	1.3	71	22	AAW38338	Peptide #12375 enc	893	6	1.3	81	16	AAAR69523	Anti-HCV antibody
821	6	1.3	71	23	ABG47122	Human peptide enco	894	6	1.3	81	16	AAAR69525	Anti-HCV antibody
822	6	1.3	71	23	ABP01417	Human ORFX protein	895	6	1.3	81	16	AAAR69528	Anti-HCV antibody
823	6	1.3	72	18	AAW28171	Amino acid sequenc	896	6	1.3	81	16	AAAR69538	Anti-HCV antibody
824	6	1.3	72	20	AAV12499	Human 5' EST secre	897	6	1.3	81	16	AAAR69533	Anti-HCV antibody
825	6	1.3	72	21	AAV65676	C. elegans insulin	898	6	1.3	81	17	AAV15430	HCV NS-3 reactive
826	6	1.3	72	21	AAO02219	Human polypeptide	899	6	1.3	81	17	AAV15431	HCV NS-3 reactive
827	6	1.3	73	21	AAV99346	Human PRO831 (UNQ4	900	6	1.3	81	17	AAV15432	HCV NS-3 reactive
828	6	1.3	73	22	AAU49580	Protonibacterium	901	6	1.3	81	17	AAV15433	HCV NS-3 reactive
829	6	1.3	73	22	ABB29349	Peptide #2000 enco	902	6	1.3	81	17	AAV15434	HCV NS-3 reactive
830	6	1.3	73	22	ABB34520	Peptide #2026 enco	903	6	1.3	81	17	AAV15435	HCV NS-3 reactive
831	6	1.3	73	22	ABB19932	Protein #1931 enco	904	6	1.3	81	17	AAV15436	HCV NS-3 reactive
832	6	1.3	73	22	AAW55307	Human brain expres	905	6	1.3	81	17	AAV15437	HCV NS-3 reactive
833	6	1.3	73	22	AAW67706	Human bone marrow	906	6	1.3	81	17	AAV15436	Protease peptide
834	6	1.3	73	22	AAW15514	Peptide #1948 enco	907	6	1.3	81	17	AAV15429	HCV NS-3 reactive
835	6	1.3	73	22	AAW24231	Human EST encoded	908	6	1.3	81	23	ABG65108	Human albumin fusi
836	6	1.3	73	22	AAW28005	Peptide #2042 enco	909	6	1.3	81	23	ABG65109	Human albumin fusi
837	6	1.3	73	22	AAW03267	Peptide #1949 enco	910	6	1.3	81	23	ABP03976	Human ORFX protein
838	6	1.3	73	22	AAW06773	Human foetal prote	911	6	1.3	81	23	ABB94023	Human secreted pro
839	6	1.3	73	22	AAAB7540	Human PRO831. Hom	912	6	1.3	81	23	ABB94047	Human secreted pro
840	6	1.3	73	22	AAAB6095	Protein of the inv	913	6	1.3	82	22	AAU64756	Protonibacterium
841	6	1.3	73	23	ABG37269	Human peptide enco	914	6	1.3	82	22	AAU67679	Human colon cancer
842	6	1.3	74	20	AAW89912	Antigen 3 from clu	915	6	1.3	83	21	AAV58542	Rat somatostatin r
843	6	1.3	74	20	ABP35133	Human ORF4106 prot	916	6	1.3	83	21	AAV58543	Mouse somatostatin
844	6	1.3	75	21	AAW54948	Arabidopsis thalia	917	6	1.3	83	21	AAV58544	Human somatostatin
845	6	1.3	75	22	AAU64148	Protonibacterium	918	6	1.3	83	22	AAU61490	Protonibacterium
846	6	1.3	75	22	ABG11271	Novel human diagno	919	6	1.3	83	22	AAW83245	Human immune/haema
847	6	1.3	75	22	AAO05523	Human polypeptide	920	6	1.3	83	22	AAO09891	Human polypeptide
848	6	1.3	75	22	AAO12191	Human polypeptide	921	6	1.3	83	22	ABG40606	Human polypeptide
849	6	1.3	76	21	AAW01738	Human secreted pro	922	6	1.3	84	16	AAAR69530	Anti-HCV antibody
850	6	1.3	76	22	AAU58988	Protonibacterium	923	6	1.3	84	17	AAV15443	HCV NS-3 reactive
851	6	1.3	76	22	ABBA1840	Peptide #9346 enco	924	6	1.3	84	22	AAU41307	Protonibacterium
852	6	1.3	76	22	AAW62713	Human brain expres	925	6	1.3	84	22	AAU67297	Protonibacterium
853	6	1.3	76	22	AAW75531	Human bone marrow	926	6	1.3	84	22	ABG21779	Novel human diagno
854	6	1.3	76	22	AAO13858	Human polypeptide	927	6	1.3	84	22	ABG24914	Novel human diagno
855	6	1.3	76	22	AAW35638	Peptide #9675 enco	928	6	1.3	84	22	AAAB51246	Mouse amino acid s
856	6	1.3	76	23	AAU91254	Nucleic acid selec	929	6	1.3	84	23	ABP30853	Streptococcus poly
857	6	1.3	77	14	AAAR34939	p270 polypeptide.	930	6	1.3	85	21	AAAG1874	Arabidopsis thalia
858	6	1.3	77	17	AAV15428	HCV NS-3 reactive	931	6	1.3	85	21	AAW52520	Arabidopsis thalia
859	6	1.3	77	17	AAAR7560	Tomato p270 cDNA-e	932	6	1.3	85	22	AAU46176	Protonibacterium
860	6	1.3	77	18	AAW08365	Tomato ovary-speci	933	6	1.3	85	22	ABG06213	Novel human diagno
861	6	1.3	77	22	ABG06203	Novel human diagno	934	6	1.3	85	22	ABB17157	Human nervous syst
862	6	1.3	77	22	ABG06427	Novel human diagno	935	6	1.3	85	22	AAW60686	Human reproductive
863	6	1.3	77	22	AAW87596	Human immune/haema	936	6	1.3	85	23	AAU81757	Partial mouse clk-
864	6	1.3	77	23	ABBA8184	Listeria monocytog	937	6	1.3	86	22	AAU56810	Protonibacterium
865	6	1.3	78	19	AAW69195	Precursor sequence	938	6	1.3	86	22	ABB15822	Human nervous syst
866	6	1.3	78	21	AAAB32697	Eucalyptus grandis	939	6	1.3	88	22	AAU49349	Protonibacterium
867	6	1.3	78	22	ABBR32134	Peptide #4785 enco	940	6	1.3	88	23	ABP34746	Human ORFX719 prot
868	6	1.3	78	22	ABBR37385	Peptide #4891 enco	941	6	1.3	89	10	AAAP92025	Sequence encoded i
869	6	1.3	78	22	ABBR2675	Protein #4674 enco	942	6	1.3	89	21	AAAG11334	Arabidopsis thalia
870	6	1.3	78	22	AAW58069	Human brain expres	943	6	1.3	89	22	ABG06199	Novel human diagno
871	6	1.3	78	22	AAW70512	Human bone marrow	944	6	1.3	89	22	ABG14499	Novel human diagno
872	6	1.3	78	22	AAW18348	Peptide #4782 enco	945	6	1.3	89	22	ABG22646	Novel human diagno
873	6	1.3	78	22	AAW25353	Human protein sequ	946	6	1.3	89	22	ABG26397	Novel human diagno
874	6	1.3	78	22	AAW30833	Peptide #4870 enco	947	6	1.3	89	22	ABBA44016	Peptide #11522 enc
875	6	1.3	78	22	AAW05955	Peptide #4637 enco	948	6	1.3	89	22	ABBR26922	Protein #8921 enco
876	6	1.3	78	23	ABGA0156	Human peptide enco	949	6	1.3	89	22	AAW65031	Human brain expres
877	6	1.3	79	15	AAAR59834	Hepatitis C virus	950	6	1.3	89	22	AAW77746	Human bone marrow
878	6	1.3	79	19	AAW75173	Human secreted pro	951	6	1.3	89	22	AAW21657	Peptide #11997 enc
879	6	1.3	79	21	AAAB58924	Breast and ovarian	952	6	1.3	89	22	AAW37960	S. plisteinaespiral
880	6	1.3	79	22	AAW92329	Human digestive sy	953	6	1.3	89	22	AAW46512	Human polypeptide
881	6	1.3	79	22	AAO13703	Human polypeptide	954	6	1.3	89	23	ABG466774	Human peptide enco
882	6	1.3	79	22	AAAG71305	Human gene 19-enco	955	6	1.3	89	23	ABBS5012	Lactococcus lactis
883	6	1.3	79	23	ABG63472	Human albumin fusi	956	6	1.3	90	15	AAAR59835	Hepatitis C virus
884	6	1.3	79	23	ABG43707	Human peptide enco	957	6	1.3	90	20	AAV38406	Human secreted pro
885	6	1.3	79	23	ABP05686	Human ORFX protein	958	6	1.3	90	21	AAAG02157	Human secreted pro
886	6	1.3	80	16	AAAR74230	HCV antigenic NS3	959	6	1.3	90	23	ABP11487	Human ORFX protein

960	6	1.3	91	18	AAW27796	Iron (III) dicitra
961	6	1.3	91	21	AAI6784	Bacteriophage Dp-1
962	6	1.3	91	21	AAI61333	Arabidopsis thalia
963	6	1.3	91	22	AAW93837	Human excretory re
964	6	1.3	91	22	ABR03894	Human musculoskele
965	6	1.3	91	22	AAW2652	Human kidney relat
966	6	1.3	92	20	AAI13124	Human secreted pro
967	6	1.3	92	21	AAI34735	Human secreted pro
968	6	1.3	92	21	AAW60105	Arabidopsis thalia
969	6	1.3	92	21	AAW69209	Amino acid sequenc
970	6	1.3	92	22	AAW63023	Propionibacterium
971	6	1.3	92	22	AAW65867	Propionibacterium
972	6	1.3	92	22	AAW31014	Novel human secret
973	6	1.3	92	22	AAW89224	Human immune/haema
974	6	1.3	93	16	AAW69529	Anti-HCV antibody
975	6	1.3	93	16	AAW69534	Anti-HCV antibody
976	6	1.3	93	22	ABG18825	Novel human diagno
977	6	1.3	94	21	AAW32851	Eucalyptus grandis
978	6	1.3	94	23	ABW78910	Tumour necrosis fa
979	6	1.3	94	23	AAW49145	Human immune/haema
980	6	1.3	95	22	AAW82228	Human immune/haema
981	6	1.3	96	20	AAW40019	Peptide sequence d
982	6	1.3	96	21	AAW34178	Zea mays protein f
983	6	1.3	96	22	ABW70684	Drosophila melanog
984	6	1.3	96	22	ABW17878	Human nervous syst
985	6	1.3	96	22	AAW02114	Human polypeptide
986	6	1.3	96	22	AAW06006	Human polypeptide
987	6	1.3	96	23	ABW08519	Human ORF protein
988	6	1.3	97	11	AAW05246	Amino acid sequenc
989	6	1.3	98	19	AAW37384	Hepatitis C virus
990	6	1.3	98	21	AAW15079	Arabidopsis thalia
991	6	1.3	98	21	AAW00199	Human secreted pro
992	6	1.3	98	23	AAW47340	Listeria monocytog
993	6	1.3	99	21	AAW60104	Arabidopsis thalia
994	6	1.3	99	21	AAW73378	HTW clone 1646005
995	6	1.3	99	23	ABW66956	Schizosaccharomyce
996	6	1.3	100	22	ABG11267	Novel human diagno
997	6	1.3	100	22	AAW04053	Human polypeptide
998	6	1.3	100	23	ABW5723	Cyclin-dependent k
999	6	1.3	101	21	AAW34177	Zea mays protein f
1000	6	1.3	101	21	AAW00200	Human secreted pro

## ALIGNMENTS

## RESULT 1

AAW05734 standard; Protein; 463 AA.

AAW05734:

23-JUL-1997 (first entry)

Presenilin-1-2.

Presenilin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;

familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;

depression; antibody; gene expression modulator; therapy.

Homo sapiens.

MO9634099-A2.

31-OCT-1996.

29-APR-1995.

31-JUL-1995.

28-APR-1995.

28-JUN-1995.

(HSCR-) HSC RES &amp; DEV LP.

PA (UTOR ) UNIV TORONTO GOVERNING COUNCIL.

XX Fraser PE, Rommens JM, St George-Hyslop PH;

XX WPT, 1996-497631/49.

DR N-PSDB: AAT40029.

XX

PT New presenilin genes - useful for diagnosis, therapy and drug

PT screening of familial Alzheimer's disease, cerebral disorders, etc.

XX

PS Claim 2; Page 132-134; 178pp; English.

XX

CC AAW05733 and AAW05734 represent the two different forms of wild type

CC human presenilin-1 (PS-1). This form of presenilin-1 results from

CC alternate splicing of the genomic DNA sequence. AAW05762 represents the

CC coding sequence for wild type human PS-2. The presenilins are a family

CC of highly conserved integral membrane proteins with a common structural

CC motif, common alternate splicing patterns, and common mutational hot

CC spot regions. Mutations in PS genes are implicated in familial

CC Alzheimer's disease (AD) and possibly other diseases such as cerebral

CC haemorrhage, schizophrenia, depression etc., so detection of mutations in

CC the DNA encoding these sequences can be used for diagnosis of these

CC diseases. These proteins, or vectors that express them or containing

CC antisense sequences, antibodies selective for mutant forms of these

CC proteins (such as AAW05736) and modulators of PS gene expression are

CC potentially useful for treatment of AD etc. Transgenic animals are useful

CC as models for drug screening. The antibodies can also be used e.g. for

XX affinity purification and in immunoassays.

XX

Sequence 463 AA:

Query Match 93.8%; Score 438; DB 17; Length 463;

Best Local Similarity 100.0%; Pred. No. 0; Matches 438; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	30	NDNREOEHNDRLSLGHPEPLSNGRPGNSRQYVEODEEDELTLTKYGAHVIMLFVPV	89
DB	26	NDNREOEHNDRLSLGHPEPLSNGRPGNSRQYVEODEEDELTLTKYGAHVIMLFVPV	85
QY	90	TLCMVVVAATIKSVSFTRKDGQILYPTFEDTETVGORLHSLINAAIMISIVVMTIL	149
DB	86	TLCMVVVAATIKSVSFTRKDGQILYPTFEDTETVGORLHSLINAAIMISIVVMTIL	145
QY	150	LVVLYKRCYKVIHAWLILSSLLLEFFSFYLGVEFKTYNVAVDYITVALLIMNLGVG	209
DB	146	LVVLYKRCYKVIHAWLILSSLLLEFFSFYLGVEFKTYNVAVDYITVALLIMNLGVG	205
QY	210	MISIHMKGPLRLOQAYILMISALMAVETIKYLPDEWAMLLAVISYDVAVLCPRGPLR	269
DB	206	MISIHMKGPLRLOQAYILMISALMAVETIKYLPDEWAMLLAVISYDVAVLCPRGPLR	265
QY	270	MLVETAOERNETLEFPALITYSTWVWLNNAGDPDEAQRVRSKSKNAESTERESODTYA	329
DB	266	MLVETAOERNETLEFPALITYSTWVWLNNAGDPDEAQRVRSKSKNAESTERESODTYA	325
QY	330	ENDDGFSEMEARQSHLQPHRSTPESRAAYQELSSILAGEDPERGVKGLGDFITY	389
DB	326	ENDDGFSEMEARQSHLQPHRSTPESRAAYQELSSILAGEDPERGVKGLGDFITY	385
QY	390	SVLVGRASATASGDWNTTACFAVAILGLCTLLLAIRKALPALPISITFGVLVFPAT	449
DB	386	SVLVGRASATASGDWNTTACFAVAILGLCTLLLAIRKALPALPISITFGVLVFPAT	445
QY	450	DYLVQPFMDQLAFHOFYI 467	
DB	446	DYLVQPFMDQLAFHOFYI 463	

## RESULT 2

AAW23965 standard; Protein; 463 AA.

AAW23965:



XX 20-JUL-1998 (first entry)  
DT  
XX  
DE Human presenilin-1.  
XX  
KW Presenilin-1; PS1 gene; human; familial Alzheimer's disease; FAD;  
KW cerebral haemorrhage; schizophrenia; depression; epilepsy;  
KW mental retardation; diagnosis; therapy; transgenic animal.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT MISC-difference 177 /note= "Phe117Ser mutation site (Claim 1)"  
FT MISC-difference 439 /note= "Ile439Val mutation site (Claim 1)"  
FT MISC-difference 253 /note= "Asp2537Ala mutation site, associated with  
FT MISC-difference 254..286 residue 254-286 deletion (Claim 1)"  
FT MISC-difference 254..286 /note= "residue 254-286 deletion mutant, associated  
FT with Asp253Ala mutation (Claim 1)"  
XX  
XX WO9801549-A2.  
PD 15-JAN-1998.  
XX  
XX 04-JUL-1997: 97WO-CA00475.  
XX  
XX 02-JAN-1997: 97US-0034590.  
PR 05-JUL-1996: 96US-0021673.  
PR 12-JUL-1996: 96US-0021700.  
PR 08-NOV-1996: 96US-0029895.  
XX  
PA (HSCR-) HSC RES & DEV LP.  
PA (UTOR ) UNIV TORONTO GOVERNING COUNCIL.  
PI Fraser PE, Rommens JM, St George-Hyslop PH;  
XX  
XX WPI: 1998-286355/25.  
DR N-PSDB: AAV04667.  
XX  
XX  
PT New isolated mutant presenilin-1 genes - useful for developing  
PT products for use in detection, diagnosis and therapy of Alzheimer's  
PT disease and for drug screening  
XX  
PS Claim 1: Page 185-186; 238pp; English.  
XX  
XX This polypeptide comprises human presenilin-1 (hps1). Its amino  
XX acid sequence was deduced from an isolated cDNA clone (see AAV04667).  
XX Another hps1 sequence (see AAW23964) results from alternative  
XX splicing of the hps1 mRNA transcript. A murine PS1 homologue (see  
XX AAW23966) and a human presenilin-2 protein (see AAW23967) are also  
XX provided. Mutations in the PS-1 and PS-2 genes are linked to the  
XX development in humans of forms of familial Alzheimer's disease  
XX (FAD) and may be causative of other disorders, e.g. cognitive,  
XX intellectual, neurological or physiological disorders such as  
XX cerebral haemorrhage, schizophrenia, depression, mental retardation  
XX and epilepsy. Use of the nucleic acids and proteins comprising or  
XX derived from the presenilins is made in screening and diagnosing  
XX FAD, identifying and developing therapeutics for treatment of FAD,  
XX and in producing cell lines and transgenic animals useful as models  
XX of FAD. Methods for identifying substances that bind to, or  
XX modulate the activity of a presenilin protein, and methods for  
XX identifying substances that affect the interaction of a  
XX presenilin-interacting protein with a presenilin protein are also  
XX disclosed.  
XX  
XX Sequence 463 AA:  
SO

QY 30 NDNREROEHNDRRSLGHPPEPLSNRPOGNSRQVVEODEEDELTLKYGAKHVMLEFPV 89  
DB 26 NDNREROEHNDRRSLGHPPEPLSNRPOGNSRQVVEODEEDELTLKYGAKHVMLEFPV 85  
QY 90 TLCAVVVVVATIKSVSFYTRKDGQILYPTFEDTEVQGRALSHILMAIMISVIVMTIL 149  
DB 86 TLCAVVVVVATIKSVSFYTRKDGQILYPTFEDTEVQGRALSHILMAIMISVIVMTIL 145  
QY 150 LVLYKRYCYVVIHAMLISSLLLFPSFYLGEPVTVVAVDYTVALLINLGVG 209  
DB 146 LVLYKRYCYVVIHAMLISSLLLFPSFYLGEPVTVVAVDYTVALLINLGVG 205  
QY 210 MISIHMKGPLRLOOAYLIMISALMALVFIKLPENTAMLIAVISYDLVAVLCPKGLR 269  
DB 206 MISIHMKGPLRLOOAYLIMISALMALVFIKLPENTAMLIAVISYDLVAVLCPKGLR 265  
QY 270 MLVETAOERNETLPPALISSTWVLMVMAEGDEPAQRVSKNSKYNNESTERESQDTVA 329  
DB 266 MLVETAOERNETLPPALISSTWVLMVMAEGDEPAQRVSKNSKYNNESTERESQDTVA 325  
QY 330 ENDGGFSEEWAEQDRSHLGRHSTPESRAVQELSSILAGEPDEEGVRLGIDFIFY 389  
DB 326 ENDGGFSEEWAEQDRSHLGRHSTPESRAVQELSSILAGEPDEEGVRLGIDFIFY 385  
QY 390 SVLVGKASATASGDMNTIACEVAIILIGLCTLLLLAIFKKALPALPISIFGLVFYPAT 449  
DB 386 SVLVGKASATASGDMNTIACEVAIILIGLCTLLLLAIFKKALPALPISIFGLVFYPAT 445  
QY 450 DYLVQPFMDQLAFHQFYI 467  
DB 446 DYLVQPFMDQLAFHQFYI 463

RESULT 3  
AAW05733  
ID AAW05733 standard; Protein; 467 AA.  
XX  
XX AAW05733:  
XX  
XX 23-JUL-1997 (first entry)  
XX  
XX Presenilin-1-1.  
XX  
XX  
XX Presenilin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;  
XX familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;  
XX depression; antibody; gene expression modulator; therapy.  
XX  
XX Homo sapiens.  
XX  
XX WO9634099-A2.  
XX  
XX 31-OCT-1996.  
XX  
XX 29-APR-1996: 96WO-CA00263.  
XX  
XX 31-JUL-1995: 95US-0509359.  
PR 28-APR-1995: 95US-0431048.  
PR 28-JUN-1995: 95US-0496841.  
XX  
XX (HSCR-) HSC RES & DEV LP.  
XX (UTOR ) UNIV TORONTO GOVERNING COUNCIL.  
XX  
XX Fraser PE, Rommens JM, St George-Hyslop PH;  
XX  
XX WPI: 1996-497631/49.  
DR N-PSDB: AAT40028.  
XX  
XX New presenilin genes - useful for diagnosis, therapy and drug  
XX screening of familial Alzheimer's disease, cerebral disorders, etc.  
XX  
XX Claim 2: Page 128-130; 178pp; English.

AAW05733 and AAW05734 represent the two different forms of wild type human presenilin-1 (PS-1). The form represented by AAW05734 results from alternate splicing of the genomic DNA sequence. AAW05762 represents the coding sequence for wild type human PS-2. The presenilins are a family of highly conserved integral membrane proteins with a common structural motif, common alternate splicing patterns, and common mutational hot spot regions. Mutations in PS genes are implicated in familial Alzheimer's disease (AD) and possibly other diseases such as cerebral haemorrhage, schizophrenia, depression etc., so detection of mutations in the DNA encoding these sequences can be used for diagnosis of these diseases. These proteins, or vectors that express them or containing antisense sequences, antibodies selective for mutant forms of these proteins (such as AAW05736) and modulators of PS gene expression are potentially useful for treatment of AD etc. Transgenic animals are useful as models for drug screening. The antibodies can also be used e.g. for affinity purification and in immunoassays.

Sequence 467 AA;

Query Match 78.4%; Score 366; DB 17; Length 467;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MTELPALSTFQNAQMSDNHLSNTVRSQDNRRQCHNDRSGHPPLSNGRPGNSR 60
DB 1 MTELPALSTFQNAQMSDNHLSNTVRSQDNRRQCHNDRSGHPPLSNGRPGNSR 60
QY 61 QVVEDEDEDEBELTKGAKHVIMLFVPTLCMVVVATIKSVSFYRKQGLYTPFTE 120
DB 61 QVVEDEDEDEBELTKGAKHVIMLFVPTLCMVVVATIKSVSFYRKQGLYTPFTE 120
QY 121 DTEVGGRALHSILNAAIMTSIVVMTLLVLYKRCYVIAHMLISLLLFPSFT 180
DB 121 DTEVGGRALHSILNAAIMTSIVVMTLLVLYKRCYVIAHMLISLLLFPSFT 180
QY 181 YLGEVFETYVAVDYITVALLINMLGVVGMSTHMKGPLKQAYLIMISALNVITKY 240
DB 181 YLGEVFETYVAVDYITVALLINMLGVVGMSTHMKGPLKQAYLIMISALNVITKY 240
QY 241 LPEWTAVLLIAVISVYVLPVLCFKGPLRMVETAOERNETLPALYSSMWLVMAE 300
DB 241 LPEWTAVLLIAVISVYVLPVLCFKGPLRMVETAOERNETLPALYSSMWLVMAE 300
QY 301 GDPDAQRRVSKNSKYNAESTERESQDTVAENDDGFSEWEAQRDSHLGRSTPESRAA 360
DB 301 GDPDAQRRVSKNSKYNAESTERESQDTVAENDDGFSEWEAQRDSHLGRSTPESRAA 360
QY 361 VOELSSSILAGEDEEENGVLGLGDFTFYSLVKGASATASGDMNTTIACFVALIGLCL 420
DB 361 VOELSSSILAGEDEEENGVLGLGDFTFYSLVKGASATASGDMNTTIACFVALIGLCL 420
QY 421 TLLLLAIFKKALPALPISITFGVLFYFATDYLVOPFMDOLAFHOFT 467
DB 421 TLLLLAIFKKALPALPISITFGVLFYFATDYLVOPFMDOLAFHOFT 467
```

RESULT 4  
ID AAW23964 standard; Protein; 467 AA.

AAW23964;

20-JUL-1998 (first entry)

Human presenilin-1.

XX Presentin-1; PS1 gene; human; familial Alzheimer's disease; FAD;  
KW cerebral haemorrhage; schizophrenia; depression; epilepsy;  
XX mental retardation; diagnosis; therapy; transgenic animal.

OS Homo sapiens.

FH Key Location/Qualifiers

```
FT Domain
FT /label= TM1
FT /note= "transmembrane domain 1"
FT Domain
FT /label= TM1-2
FT /note= "hydrophilic loop"
FT Domain
FT /label= TM2
FT /note= "transmembrane domain 2"
FT Domain
FT /label= TM2-3
FT /note= "hydrophilic loop"
FT Domain
FT /label= TM3
FT /note= "transmembrane domain 3"
FT Domain
FT /label= TM3-4
FT /note= "hydrophilic loop"
FT Domain
FT /label= TM4
FT /note= "transmembrane domain 4"
FT Domain
FT /label= TM4-5
FT /note= "hydrophilic loop"
FT Domain
FT /label= TM5
FT /note= "transmembrane domain 5"
FT Domain
FT /label= TM5-6
FT /note= "hydrophilic loop"
FT Domain
FT /label= TM6
FT /note= "transmembrane domain 6"
FT Domain
FT /label= TM6-7
FT /note= "hydrophilic loop"
FT Domain
FT /label= TM8
FT /note= "transmembrane domain 8"
FT Misc-difference
FT /note= "Phe117Ser mutation site (Claim 1)"
FT Misc-difference
FT /note= "Ile439Val mutation site (Claim 1)"
FT Misc-difference
FT /note= "Asp257Ala mutation site, associated with residue 258-290 deletion (Claim 1)"
FT Misc-difference
FT /note= "residue 258-290 deletion mutant, associated with Asp257Ala mutation (Claim 1)"
FT Misc-difference
FT /note= "Ile143Thr mutation site (Claim 18)"
FT Misc-difference
FT /note= "Met146Leu mutation site (Claim 18)"
FT Misc-difference
FT /note= "Leu171Pro mutation site (Claim 18)"
FT Misc-difference
FT /note= "Ala260Val mutation site (Claim 18)"
FT Misc-difference
FT /note= "Cys263Arg mutation site (Claim 18)"
FT Misc-difference
FT /note= "Pro264Leu mutation site (Claim 18)"
FT Misc-difference
FT /note= "Pro267Ser mutation site (Claim 18)"
FT Misc-difference
FT /note= "Glu280Ala mutation site (Claim 18)"
FT Misc-difference
FT /note= "Glu280Gly mutation site (Claim 18)"
FT Misc-difference
FT /note= "Ala285Val mutation site (Claim 18)"
FT Misc-difference
FT /note= "Leu286Val mutation site (Claim 18)"
FT Misc-difference
FT /note= "Leu322Val mutation site (Claim 18)"
```

FT	Misc-difference	392	
FT	/note-	"Leu32Val mutation site (Claim 18)"	
FT	Misc-difference	410	
FT	/note-	"Cys410Tyr mutation site (Claim 18)"	
FT	Misc-difference	79	
FT	/note-	"Ala79Xaa mutation site"	
FT	Misc-difference	82	
FT	/note-	"Val82Leu mutation site"	
FT	Misc-difference	96	
FT	/note-	"Val96Phe mutation site"	
FT	Misc-difference	115	
FT	/note-	"Tyr115His mutation site"	
FT	Misc-difference	139	
FT	/note-	"Met139Thr mutation site"	
FT	Misc-difference	139	
FT	/note-	"Met139Val mutation site"	
FT	Misc-difference	146	
FT	/note-	"Met146Val mutation site"	
FT	Misc-difference	163	
FT	/note-	"His163Arg mutation site"	
FT	Misc-difference	163	
FT	/note-	"His163Tyr mutation site"	
FT	Misc-difference	209	
FT	/note-	"Gly209Val mutation site"	
FT	Misc-difference	211	
FT	/note-	"Ile211Thr mutation site"	
FT	Misc-difference	231	
FT	/note-	"Ala231Thr mutation site"	
FT	Misc-difference	246	
FT	/note-	"Ala246Glu mutation site"	
FT	Misc-difference	291..319	
FT	/note-	"residue 291-319 deletion site"	
FT	Misc-difference	384	
FT	/note-	"Gly384Ala mutation site"	
FT	Misc-difference	26..29	
FT	/note-	"residue 26-29 deletion, resulting from alternative splicing"	
PN	WO9801549-A2.		
PD	15-JAN-1998.		
XX			
PF	04-JUL-1997:	97MO-CA00475.	
XX			
PR	02-JAN-1997:	97US-0034590.	
PR	05-JUL-1996:	96US-0021673.	
PR	12-JUL-1996:	96US-0021700.	
PR	08-NOV-1996:	96US-0028855.	
XX	(HSCR-) HSC RES & DEV LP.		
PA	(UTOR ) UNIV TORONTO GOVERNING COUNCIL.		
PI	Fraser PE, Rommens JM, St George-Hyslop PH;		
XX			
DR	WPI: 1998-286355/25.		
XX	N-PSDB: AAV04666.		
PT	New isolated mutant presenilin-1 genes - useful for developing		
PT	products for use in detection, diagnosis and therapy of Alzheimer's		
PT	disease and for drug screening		
XX			
PS	Claim 1: Page 180-182; 238pp; English.		
CC	This polypeptide comprises human, presenilin-1 (hps1). Its amino		
CC	acid sequence was deduced from an isolated cDNA clone (see AAV04666)		
CC	Another hps1 sequence (see AAW23965) results from alternative		
CC	splicing of the hps1 mRNA transcript. A murine PSI homologue (see		
CC	AAW23966) and a human presenilin-2 protein (see AAW23967) are also		
CC	provided. Mutations in the PS-1 and PS-2 genes are linked to the		
CC	development in humans of forms of familial Alzheimer's disease		
CC	(FAD) and may be causative of other disorders, e.g. cognitive,		
CC	intellectual, neurological or physiological disorders such as		
CC	cerebral haemorrhage, schizophrenia, depression, mental retardation		

	Query Match	Best Local Similarity	Matches 466; Conservative	Score 366; DB 19; Length 467; 78.4%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 MTELPAPLSTYFONAAQMSSEDNHLSNVRSONDRREOEHNDRSLGHPPLSGRPOGNSR	60		
Db	1 MTELPAPLSTYFONAAQMSSEDNHLSNVRSONDRREOEHNDRSLGHPPLSGRPOGNSR	60		
QY	61 QVEODEDEDEBELTLKYGAKHVIMLFVPTLCMVVVVATIKSFYTRKDGOLITYPTE	120		
Db	61 QVEODEDEDEBELTLKYGAKHVIMLFVPTLCMVVVVATIKSFYTRKDGOLITYPTE	120		
QY	121 DTEYGGORALHSILNAAIMISYIVMTTLLVLYTRCYKVIHAMIISLLLPFFSFI	180		
Db	121 DTEYGGORALHSILNAAIMISYIVMTTLLVLYTRCYKVIHAMIISLLLPFFSFI	180		
QY	181 YLGEVFKTYNNAVDITYTALLIMNLGVGMISIHMKGPRLQOAVLIMISALMALVFITY	240		
Db	181 YLGEVFKTYNNAVDITYTALLIMNLGVGMISIHMKGPRLQOAVLIMISALMALVFITY	240		
QY	241 LPEWTAMILIAVISYVDLVAVLCPRKPLRLMETAEQERNETLFPALISSTWMLVNNAE	300		
Db	241 LPEWTAMILIAVISYVDLVAVLCPRKPLRLMETAEQERNETLFPALISSTWMLVNNAE	300		
QY	301 GDPKQRRYSKSKSTNAESTERESODTYAENDDGFSEBEKQORSHLCPHRSTPESRAA	360		
Db	301 GDPKQRRYSKSKSTNAESTERESODTYAENDDGFSEBEKQORSHLCPHRSTPESRAA	360		
QY	361 VOELSSSLIAGDPEERGVKLGIDPIFYSVLVGRASATASGDMWTTTACFALLIGLCL	420		
Db	361 VOELSSSLIAGDPEERGVKLGIDPIFYSVLVGRASATASGDMWTTTACFALLIGLCL	420		
QY	421 TLLLLAIFKKALPALPISITTEGLVVFATDYLQVFPMDLAFHQEYI	467		
Db	421 TLLLLAIFKKALPALPISITTEGLVVFATDYLQVFPMDLAFHQEYI	467		
RESULT 5				
AAW41430				
ID	AAW41430 standard; Protein: 467 AA.			
XX	AAW41430;			
AC				
XX	04-JUN-1998 (first entry)			
DT				
XX	PS1/467 protein.			
DE				
XX	Presentin peptide; PS1/429; immunogen; immune response; PS1 gene;			
KW	Alzheimer's disease; mitochondrial pathology; neurodegeneration;			
KW	apoptosis; PS1/467.			
XX				
OS	Homo sapiens.			
XX				
PN	W09746678.1.			
XX				
PD	11-DEC-1997.			
XX				
PF	03-JUN-1997; 9AWO-US09272.			
XX				
XX	18-JUL-1996; 96US-0683315.			
PR	06-JUN-1996; 96US-0659296.			



```

Db 181 YLGEVFKTYNNAVDTITVALIIMNFGVGMISIHMKGPLRLQOAYLIMISALMALVFIKY 240
Oy 241 LPETWAMILIAVISYDVAVALCPKGPLRLMLVETAOERRETLFPALISSTWVWLNNAE 300
Db 241 LPETWAMILIAVISYDVAVALCPKGPLRLMLVETAOERRETLFPALISSTWVWLNNAE 300
Oy 301 GDPENQRRVSKSKYNAESTERESQDTVAENDDGFSEEMEAQRSHLGPHRSTPESRAA 360
Db 301 GDPENQRRVSKSKYNAESTERESQDTVAENDDGFSEEMEAQRSHLGPHRSTPESRAA 360
Oy 361 VOELSSSTILAGEDEPERGVKLGIDGFIYSVLVGKASATASGDMWNTTACFVAAILIGLCL 420
Db 361 VOELSSSTILAGEDEPERGVKLGIDGFIYSVLVGKASATASGDMWNTTACFVAAILIGLCL 420
Oy 421 TLLLAIFKKALPALPISITFGLVYFANDYLVOPFMOLAHQFYI 467
Db 421 TLLLAIFKKALPALPISITFGLVYFANDYLVOPFMOLAHQFYI 467

RESULT 7
AAY23897
ID AAY23897 standard; protein: 467 AA.
AC AAY23897;
XX 27-SEP-1999 (first entry)
DT 27-SEP-1999 (first entry)
DE Amino acid sequence of the wild type human presenilin 1 (PS1) protein.
XX Human; presenilin 1; PS1; presenilin-binding protein; interacting domain;
KW presenilin allele; Alzheimer's disease; senile dementia;
KW psychiatric disease; schizophrenia; depression; neurological disease;
KW stroke; cerebral haemorrhage.
XX Homo sapiens.
OS Homo sapiens.
XX MO9935501-A1.
XX 15-JUL-1999.
PD 15-JUL-1999.
XX 08-JAN-1999; 99WO-CA00018.
PE 08-JAN-1999;
XX 09-JAN-1998; 98US-0070948.
PR 09-JAN-1998;
XX (UNIV ) UNIV TORONTO GOVERNING COUNCIL.
PA (UNIV ) UNIV TORONTO GOVERNING COUNCIL.
XX Fraser PE, St George-Hyslop PH;
PI Fraser PE, St George-Hyslop PH;
XX WPI; 1999-419410/35.
DR Identifying substances that alter presenilin interactions, useful
XX for screening individuals for presenilin alleles associated with
PT Alzheimer's disease - useful for diagnosis of Alzheimer's disease
PS Alzheimer's disease - useful for diagnosis of Alzheimer's disease
XX Disclosure: Page 34-36; 40pp; English.
XX The present sequence represents wild type human presenilin 1 (PS1)
CC protein. The specification describes a method for identifying substances
CC that alter the interaction of a presenilin with a presenilin-binding
CC protein. The method comprises contacting the interacting domain of a
CC presenilin protein to a presenilin-binding protein in the presence of
CC a test substance, and measuring the interaction of the presenilin and
CC the presenilin-binding protein. The method can be used to screen
CC individuals for presenilin alleles associated with Alzheimer's disease
CC and related disorders, such as senile dementia's, psychiatric diseases
CC such as schizophrenia and depression, and neurological disease, such as
CC stroke and cerebral haemorrhage.
XX Sequence 467 AA:
SQ

```

```

Oy 1 MTELPAPLSTYFONAMSEDNHLSTNVRSONDREROENDRSLGHPPLSGRPOGNSR 60
Db 1 MTELPAPLSTYFONAMSEDNHLSTNVRSONDREROENDRSLGHPPLSGRPOGNSR 60
Oy 61 QVEQDEEDELTLKYGAKHVMLEFVPTLCMNVVVTIKSVSFYTRKDGQILTYPFTE 120
Db 61 QVEQDEEDELTLKYGAKHVMLEFVPTLCMNVVVTIKSVSFYTRKDGQILTYPFTE 120
Oy 121 DTEYGORALHSILNAAMISVYVMTLLVLYRYCKYVHAMLISSLLLEFFSFI 180
Db 121 DTEYGORALHSILNAAMISVYVMTLLVLYRYCKYVHAMLISSLLLEFFSFI 180
Oy 181 YLGEVFKTYNNAVDTITVALIIMNFGVGMISIHMKGPLRLQOAYLIMISALMALVFIKY 240
Db 181 YLGEVFKTYNNAVDTITVALIIMNFGVGMISIHMKGPLRLQOAYLIMISALMALVFIKY 240
Oy 241 LPETWAMILIAVISYDVAVALCPKGPLRLMLVETAOERRETLFPALISSTWVWLNNAE 300
Db 241 LPETWAMILIAVISYDVAVALCPKGPLRLMLVETAOERRETLFPALISSTWVWLNNAE 300
Oy 301 GDPENQRRVSKSKYNAESTERESQDTVAENDDGFSEEMEAQRSHLGPHRSTPESRAA 360
Db 301 GDPENQRRVSKSKYNAESTERESQDTVAENDDGFSEEMEAQRSHLGPHRSTPESRAA 360
Oy 361 VOELSSSTILAGEDEPERGVKLGIDGFIYSVLVGKASATASGDMWNTTACFVAAILIGLCL 420
Db 361 VOELSSSTILAGEDEPERGVKLGIDGFIYSVLVGKASATASGDMWNTTACFVAAILIGLCL 420
Oy 421 TLLLAIFKKALPALPISITFGLVYFANDYLVOPFMOLAHQFYI 467
Db 421 TLLLAIFKKALPALPISITFGLVYFANDYLVOPFMOLAHQFYI 467

RESULT 8
AAB07971
ID AAB07971 standard; protein: 467 AA.
AC AAB07971;
XX 14-NOV-2000 (first entry)
DT 14-NOV-2000 (first entry)
DE Amino acid sequence of human presenilin I polypeptide.
XX Human; presenilin; PS1; PS2; nerve cell growth; nerve damage; hNPRAP;
KW neural plakophilin related armidillo protein; Alzheimer's disease;
KW Parkinson's disease; amyotrophic lateral sclerosis; ALS; stroke;
KW multiple sclerosis; ischemia; stroke; neural paropathy; sciatic crush;
KW motor neuron disease; peripheral neuropathy; neuropathy; diabetes;
KW spinal cord injury; facial nerve crush.
XX Homo sapiens.
OS Homo sapiens.
XX WO200047615-A2.
XX 17-AUG-2000.
PD 17-AUG-2000.
XX 11-FEB-2000; 2000WO-CA00126.
PE 11-FEB-2000;
XX 12-FEB-1998; 99US-0119835.
PR 12-FEB-1998;
XX (UNIV ) UNIV TORONTO GOVERNING COUNCIL.
PA (UNIV ) UNIV TORONTO GOVERNING COUNCIL.
XX St George-Hyslop PH, Fraser PE;
PI St George-Hyslop PH, Fraser PE;
XX WPI; 2000-524531/47.
DR Stimulation of nerve cell growth using human Neural Plakophilin Related
XX Armidillo Protein (hNPRAP) polypeptide, useful for the treatment of
PT diseases such as Alzheimer's, Parkinson's, and stroke -
PS Disclosure: Page 19-20; 33pp; English.
XX

```

CC The present sequence represents a human presenilin 1 (PS1) polypeptide.  
CC Human Neural Plakophilin Related Armadillo protein (hNPRAP) polypeptide  
CC is known to interact with PSI and PS2. The specification describes a  
CC method for stimulating the growth of nerve cells, comprising contacting  
CC them with hNPRAP. The hNPRAP polypeptide and polynucleotide are useful  
CC for treating nerve damage caused by a variety of diseases or physical  
CC traumas, including Alzheimer's disease, Parkinson's disease, amyotrophic  
CC lateral sclerosis (ALS), multiple sclerosis, stroke, ischemia associated  
CC with stroke, neural paropathy, motor neuron diseases, sciatic crush,  
CC peripheral neuropathy, neuropathy associated with diabetes, spinal cord  
CC injuries and facial nerve crush.

XX Sequence 467 AA:

Query Match 78.4%; Score 366; DB 21; Length 467;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTELPAPLSTYQNAQMSQNDNLSTVRSQNDNRROEHNDRRSLGHEPLSNGRPQNSR 60  
DB 1 MTELPAPLSTYQNAQMSQNDNLSTVRSQNDNRROEHNDRRSLGHEPLSNGRPQNSR 60  
QY 61 QVDEDEDEDELTKYGAHVIMLEFVPTLCMVVVAATIKSVSFYRKDGQLYTPPTE 120  
DB 61 QVDEDEDEDELTKYGAHVIMLEFVPTLCMVVVAATIKSVSFYRKDGQLYTPPTE 120  
QY 121 DTEVVGORALHSILNAIMISVIYVMTLLVLYKYRCYKVIHAMLIISSLLFFFSFI 180  
DB 121 DTEVVGORALHSILNAIMISVIYVMTLLVLYKYRCYKVIHAMLIISSLLFFFSFI 180  
QY 181 YLGEVFTYNAVDYITVALLIMNLGVVGMISIMWGPRLQQAYLIMISALMALVFIKY 240  
DB 181 YLGEVFTYNAVDYITVALLIMNLGVVGMISIMWGPRLQQAYLIMISALMALVFIKY 240  
QY 241 LPEWTAMLILAVISVYDVAVLCPKGPLRMLVETAOERNETLFPALIISSSTMVWLVMAE 300  
DB 241 LPEWTAMLILAVISVYDVAVLCPKGPLRMLVETAOERNETLFPALIISSSTMVWLVMAE 300  
QY 301 GDPBAQRVRSKNSKYNAESTERESQDTVAENDDGFSEMEAOQSDSHLGPARRSTPESRAA 360  
DB 301 GDPBAQRVRSKNSKYNAESTERESQDTVAENDDGFSEMEAOQSDSHLGPARRSTPESRAA 360  
QY 361 VOELSSSILAGEDPEERGVKLGDFTFYSVLVKGASATASGDMNTTIACFVAIIIGLCL 420  
DB 361 VOELSSSILAGEDPEERGVKLGDFTFYSVLVKGASATASGDMNTTIACFVAIIIGLCL 420  
QY 421 TLLLLAIFKKALPALPISITFGVFPATDYLVOFPMDQLAFHOFTY 467  
DB 421 TLLLLAIFKKALPALPISITFGVFPATDYLVOFPMDQLAFHOFTY 467

RESULT 9

ID AAE10798 standard; Protein; 467 AA.

XX AAE10798;

DT 18-DEC-2001 (first entry)

DE Human presenilin-1 (PS-1) protein.

XX Human; catenin p120; presenilin-1; PS-1; neuroprotective; gene therapy;  
KW neurodegenerative disease; Alzheimer's disease; nootropic; prophylaxis;  
KW neuronal disorder; cognitive disorder.

XX Homo sapiens.

XX WO200167097-A2.

XX 13-SEP-2001.

XX 09-MAR-2001; 2001WO-GB01059.

PR 10-MAR-2000; 2000GB-0005895.

XX (GLAXO ) GLAXO GROUP LTD.

XX Hale RS, Rowley A, Blackstock W;

XX WPI; 2001-569954/66.

DR N-PSDB; AAD18120.

PT Identifying presenilin or catenin p120 activity modulator useful for  
PT modulating presenilin-catenin p120 interaction and thus for treating  
PT cognitive disorder e.g., Alzheimer's disease comprises enhancing  
PT cognitive function

XX Example 1; Page 41-42; 48pp; English.

CC The invention relates to a method for identifying modulators of  
CC presenilin and catenin p120. Modulators of catenin p120 and presenilin  
CC are useful for the treatment and prophylaxis of disorders that is  
CC responsive to modulation of presenilin/catenin p120 activity. In  
CC particular, neuronal disorders such as cognitive disorders and  
CC neurodegenerative diseases such as Alzheimer's disease. Catenin p120 DNAs  
CC are useful for identifying mutations in catenin p120 genes.  
CC Identification of such mutations assist in the diagnosis of or  
CC susceptibility to Alzheimer's or other conditions associated with  
CC presenilin and in assessing the physiology of such disorders. Catenin  
CC p120 DNAs are also used in hybridisation studies to monitor expression of  
CC expression. The present sequence is human presenilin-1 (PS-1) protein.

XX Sequence 467 AA:

Query Match 78.4%; Score 366; DB 22; Length 467;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTELPAPLSTYQNAQMSQNDNLSTVRSQNDNRROEHNDRRSLGHEPLSNGRPQNSR 60  
DB 1 MTELPAPLSTYQNAQMSQNDNLSTVRSQNDNRROEHNDRRSLGHEPLSNGRPQNSR 60  
QY 61 QVDEDEDEDELTKYGAHVIMLEFVPTLCMVVVAATIKSVSFYRKDGQLYTPPTE 120  
DB 61 QVDEDEDEDELTKYGAHVIMLEFVPTLCMVVVAATIKSVSFYRKDGQLYTPPTE 120  
QY 121 DTEVVGORALHSILNAIMISVIYVMTLLVLYKYRCYKVIHAMLIISSLLFFFSFI 180  
DB 121 DTEVVGORALHSILNAIMISVIYVMTLLVLYKYRCYKVIHAMLIISSLLFFFSFI 180  
QY 181 YLGEVFTYNAVDYITVALLIMNLGVVGMISIMWGPRLQQAYLIMISALMALVFIKY 240  
DB 181 YLGEVFTYNAVDYITVALLIMNLGVVGMISIMWGPRLQQAYLIMISALMALVFIKY 240  
QY 241 LPEWTAMLILAVISVYDVAVLCPKGPLRMLVETAOERNETLFPALIISSSTMVWLVMAE 300  
DB 241 LPEWTAMLILAVISVYDVAVLCPKGPLRMLVETAOERNETLFPALIISSSTMVWLVMAE 300  
QY 301 GDPBAQRVRSKNSKYNAESTERESQDTVAENDDGFSEMEAOQSDSHLGPARRSTPESRAA 360  
DB 301 GDPBAQRVRSKNSKYNAESTERESQDTVAENDDGFSEMEAOQSDSHLGPARRSTPESRAA 360  
QY 361 VOELSSSILAGEDPEERGVKLGDFTFYSVLVKGASATASGDMNTTIACFVAIIIGLCL 420  
DB 361 VOELSSSILAGEDPEERGVKLGDFTFYSVLVKGASATASGDMNTTIACFVAIIIGLCL 420  
QY 421 TLLLLAIFKKALPALPISITFGVFPATDYLVOFPMDQLAFHOFTY 467  
DB 421 TLLLLAIFKKALPALPISITFGVFPATDYLVOFPMDQLAFHOFTY 467

RESULT 10

ID AAG63936 standard; Protein; 467 AA.

XX

AC AAG63936;  
 XX 29-OCT-2001 (first entry)  
 DT  
 XX Amino acid sequence of human presenilin 1.  
 DE  
 XX KIAA0253; presenilin; Alzheimer's disease.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200167109-A1.  
 XX  
 XX 13-SEP-2001.  
 XX  
 XX 09-MAR-2001; 2001WO-GB01057.  
 XX  
 XX 10-MAR-2000; 2000GB-0005894.  
 XX  
 XX (GLAX ) GLAXO GROUP LTD.  
 XX  
 XX Hale RS, Rowley A, Blackstock W;  
 XX  
 XX WPI: 2001-522960/57.  
 DR  
 XX N-PSDB: AAH74993.  
 DR  
 XX  
 XX Identifying a modulator of presenilin function by determining the  
 PT ability of presenilin to bind to a KIAA0253 polypeptide in the presence  
 PT and absence of a test compound, useful in the treatment or prophylaxis  
 PT of Alzheimer's disease -  
 PT  
 XX  
 XX Disclosure: Page 40-41; 48pp; English.  
 PS  
 XX  
 XX The present sequence represents human presenilin 1. KIAA0253 binds to  
 CC presenilin. The specification describes a method of identifying a  
 CC modulator of presenilin function or KIAA0253 function. The method  
 CC comprises determining presenilin activity or KIAA0253 activity  
 CC in the presence and absence of a test compound, where presenilin  
 CC activity is determined by its ability to bind to KIAA0253. A modulator  
 CC of presenilin or KIAA0253 polypeptide is useful in the manufacture of  
 CC a medicament for the treatment or prophylaxis of Alzheimer's disease.  
 CC The KIAA0253 polynucleotide and KIAA0253 polypeptide are useful in the  
 CC treatment, prophylaxis or diagnosis of Alzheimer's disease.  
 CC  
 XX  
 XX Sequence 467 AA:

Query Match 78.4%; Score 366; DB 22; Length 467;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MTELPAPISTYFQNAOMSEDNHLSNTVRSONDREROHNDRLSLGHPPLSGRPGNSR 60  
 DB 1 MTELPAPISTYFQNAOMSEDNHLSNTVRSONDREROHNDRLSLGHPPLSGRPGNSR 60  
 OY 61 OVVEDODEDEDELTKYGAHVIMLFVPTLCMVVVAVATIKSVSYTRKDGOLITPTPE 120  
 DB 61 OVVEDODEDEDELTKYGAHVIMLFVPTLCMVVVAVATIKSVSYTRKDGOLITPTPE 120  
 OY 121 DTEYTGQRALSHILNAAMISIVVMTILVLYKRCYKVAHMLIISLLFFFSFI 180  
 DB 121 DTEYTGQRALSHILNAAMISIVVMTILVLYKRCYKVAHMLIISLLFFFSFI 180  
 OY 181 YGGEVFKTVNAVADITVALLIMNGVGMISIHMKGPLRLOOAVLIMISALMAVFIKY 240  
 DB 181 YGGEVFKTVNAVADITVALLIMNGVGMISIHMKGPLRLOOAVLIMISALMAVFIKY 240  
 OY 241 LPEWTAMLILAVISYVDLAVLCPRGPLRMLVETAQERNETLFPALIVSSTVMVLMNAE 300  
 DB 241 LPEWTAMLILAVISYVDLAVLCPRGPLRMLVETAQERNETLFPALIVSSTVMVLMNAE 300  
 OY 301 GPPEAQRVRSKSKYNAESTERESODTVAENDGCFSEMEQRDSHLCGRHSTPESRAA 360  
 DB 301 GPPEAQRVRSKSKYNAESTERESODTVAENDGCFSEMEQRDSHLCGRHSTPESRAA 360

OY 361 VOELSSSILAGEDPEERGVKLGDFIFYSVLVGKASATASGDMWTTIACFAILIGLCL 420  
 DB 361 VOELSSSILAGEDPEERGVKLGDFIFYSVLVGKASATASGDMWTTIACFAILIGLCL 420  
 OY 421 TLLILAIFKKALPALPISTFGLVFYFATDYLVPFMDQAFHQFYI 467  
 DB 421 TLLILAIFKKALPALPISTFGLVFYFATDYLVPFMDQAFHQFYI 467

## RESULT 11

AAE05466  
 ID AAE05466 standard; Protein; 467 AA.

AC AAE05466;

DT 24-SEP-2001 (first entry)

DE Human presenilin (PS1) protein.

Human: Par-4; presenilin; PS1; neuroprotective; nuclear factor kappa B;  
 NF-kappa B; neuronal degeneration; spinal muscular atrophy; paralysis;  
 peripheral neuropathy; motorneuron disorder; neurodegenerative disorder;  
 Parkinson's disease; Meniere's disease; multiple sclerosis; Bell's palsy;  
 Huntington's chorea; Down's syndrome; amyotrophic lateral sclerosis; ALS;  
 nerve deafness; Alzheimer's disease; epilepsy.

Homo sapiens.

WO200151671-A2.

19-JUL-2001.

08-JAN-2001; 2001WO-0S00526.

10-JAN-2000; 2000US-0175200.

04-JAN-2001; 2001US-0754949.

(SCTO-) SCTOS INC.

Mccarthy J, Cordell B;

WPI: 2001-451872/48.

N-PSDB: AAE05466.

Identifying inhibitors of neuronal degeneration useful for treating  
 e.g. Alzheimer's disease, by determining the ability of a compound to  
 induce nuclear factor kappa B activation, with the involvement of  
 presenilin or Par-4 -

Claim 4; Page 60-61; 66pp; English.

The invention relates to human Par-4 protein, presenilin protein (PS1  
 and PS2) and their corresponding DNA molecules. The invention also  
 relates to a method for identifying inhibitors of neuronal degeneration,  
 comprising cotransfecting eukaryotic host cells expressing presenilin  
 (PS), with a Par-4 DNA, and an NF-kappa B dependent reporter construct,  
 exposing the cotransfected cells to a candidate molecule and monitoring  
 the ability of the candidate molecule to induce NF-kappa B activation.  
 Presenilin proteins participate in nuclear factor kappa B (NF-kappa B)  
 signaling and activation. The inhibitors of neuronal degeneration  
 are useful for treating neurodegenerative disorders such as Alzheimer's  
 disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's  
 chorea, Down's syndrome, nerve deafness, Meniere's disease and also for  
 treating peripheral neuropathies, motorneuron disorders such as  
 amyotrophic lateral sclerosis (ALS), Bell's palsy and various conditions  
 involving spinal muscular atrophy and paralysis. The present sequence  
 is human presenilin (PS1) protein.

Query Match 78.4%; Score 366; DB 22; Length 467;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 MTELPAPLSTYFONAOQMSDNHLSNTVRSQNDNRREROEHNDRRSLGHPPLSNGRPOGNSR 60
DB 1 MTELPAPLSTYFONAOQMSDNHLSNTVRSQNDNRREROEHNDRRSLGHPPLSNGRPOGNSR 60
OY 61 QVVEDDEDEDELTKYGAHVIMLFVPTLCMVVVAATIKSVSFYTRKDGQLIYTPETE 120
DB 61 QVVEDDEDEDELTKYGAHVIMLFVPTLCMVVVAATIKSVSFYTRKDGQLIYTPETE 120
OY 121 DTEYVGORALSHILNAAIMISIVVMTLLVLYKRYCKYIHAMLISSLLFFESFI 180
DB 121 DTEYVGORALSHILNAAIMISIVVMTLLVLYKRYCKYIHAMLISSLLFFESFI 180
OY 181 YLGEVFKTYNNAVDYITVALLIWNLGVGVMISIHMKGPLRLQOAYLIMISALMALVFIKY 240
DB 181 YLGEVFKTYNNAVDYITVALLIWNLGVGVMISIHMKGPLRLQOAYLIMISALMALVFIKY 240
OY 241 LPEWTAMILIAVISYVDIAVAVLCRPGPLRMIVETAOENETLFPALIIYSTWMLVYNMAE 300
DB 241 LPEWTAMILIAVISYVDIAVAVLCRPGPLRMIVETAOENETLFPALIIYSTWMLVYNMAE 300
OY 301 GDPFAORRVSKNSKYNAESTERESQDTVAENDDGFSEMEQAQRDSHILGPHRSTPESRAA 360
DB 301 GDPFAORRVSKNSKYNAESTERESQDTVAENDDGFSEMEQAQRDSHILGPHRSTPESRAA 360
OY 361 VOELSSSILAGEDPEERGVKLGIDFIYSVLVGKASATASGDMNTTACFVAALLIGLCL 420
DB 361 VOELSSSILAGEDPEERGVKLGIDFIYSVLVGKASATASGDMNTTACFVAALLIGLCL 420
OY 421 TLLLAIFKKALPALPISITFGVLFYFATDYLVOFMDQLAFHOFYI 467
DB 421 TLLLAIFKKALPALPISITFGVLFYFATDYLVOFMDQLAFHOFYI 467

RESULT 12
AA018049
ID AA018049 standard; Protein: 467 AA.
AC AA018049:
XX 02-SEP-2002 (first entry)
DE Presenilin protein.
XX Sel-12; presenilin; neuronal disorder; familial Alzheimer's disease;
KW amyloid precursor protein; APP.
XX Unidentified.
OS
XX US6376239-B1.
PN 23-APR-2002.
XX 04-APR-1997; 97US-0832867.
PF 04-APR-1997; 97US-0832867.
PR 04-APR-1997; 97US-0832867.
XX (ELEG-) ELGENE GMBH.
XX Baumeister R;
XX WPI; 2002-478281/51.
DR N-PSDB; AAL47323.
XX Isolated DNA molecule comprising promoter of the sel-12 gene from
PT Caenorhabditis elegans operably linked to heterologous gene, directs
PT expression in neural cells and is useful to develop drugs to treat
PT neuronal disorders.
XX Claim 2; Fig 4; 78pp; English.
XX The present invention relates to DNA molecules comprising the promoter of
CC the sel-12 gene from Caenorhabditis elegans operably linked to a

```

```

CC heterologous DNA sequence encoding a protein of interest. The sequence
CC can be used to develop drugs for the treatment, prevention or delay of a
CC neuronal disorder. In particular, the neuronal disorder may be familial
CC Alzheimer's disease. The present sequence is a presenilin protein
CC described in the exemplification of the invention.
XX
S0 Sequence 467 AA;
Query Match 78.4%; Score 366; DB 23; Length 467;
Best Local Similarity 99.8%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;
Matches 466; Conservative 0;
OY 1 MTELPAPLSTYFONAOQMSDNHLSNTVRSQNDNRREROEHNDRRSLGHPPLSNGRPOGNSR 60
DB 1 MTELPAPLSTYFONAOQMSDNHLSNTVRSQNDNRREROEHNDRRSLGHPPLSNGRPOGNSR 60
OY 61 QVVEDDEDEDELTKYGAHVIMLFVPTLCMVVVAATIKSVSFYTRKDGQLIYTPETE 120
DB 61 QVVEDDEDEDELTKYGAHVIMLFVPTLCMVVVAATIKSVSFYTRKDGQLIYTPETE 120
OY 121 DTEYVGORALSHILNAAIMISIVVMTLLVLYKRYCKYIHAMLISSLLFFESFI 180
DB 121 DTEYVGORALSHILNAAIMISIVVMTLLVLYKRYCKYIHAMLISSLLFFESFI 180
OY 181 YLGEVFKTYNNAVDYITVALLIWNLGVGVMISIHMKGPLRLQOAYLIMISALMALVFIKY 240
DB 181 YLGEVFKTYNNAVDYITVALLIWNLGVGVMISIHMKGPLRLQOAYLIMISALMALVFIKY 240
OY 241 LPEWTAMILIAVISYVDIAVAVLCRPGPLRMIVETAOENETLFPALIIYSTWMLVYNMAE 300
DB 241 LPEWTAMILIAVISYVDIAVAVLCRPGPLRMIVETAOENETLFPALIIYSTWMLVYNMAE 300
OY 301 GDPFAORRVSKNSKYNAESTERESQDTVAENDDGFSEMEQAQRDSHILGPHRSTPESRAA 360
DB 301 GDPFAORRVSKNSKYNAESTERESQDTVAENDDGFSEMEQAQRDSHILGPHRSTPESRAA 360
OY 361 VOELSSSILAGEDPEERGVKLGIDFIYSVLVGKASATASGDMNTTACFVAALLIGLCL 420
DB 361 VOELSSSILAGEDPEERGVKLGIDFIYSVLVGKASATASGDMNTTACFVAALLIGLCL 420
OY 421 TLLLAIFKKALPALPISITFGVLFYFATDYLVOFMDQLAFHOFYI 467
DB 421 TLLLAIFKKALPALPISITFGVLFYFATDYLVOFMDQLAFHOFYI 467

RESULT 13
AA079416
ID AA079416 standard; protein: 467 AA.
AC AA079416:
XX 02-JUL-2002 (first entry)
DE Human presenilin-1, PS-1.
XX Human; integrin-linked kinase; ILK; presenilin-1; protein kinase B; PKB;
KW gamma secretase; apoptosis; Alzheimer's disease; PS-1; neurotropic;
KW neuroprotective; cytosstatic; cancer.
XX Homo sapiens.
OS
XX WO200222862-A2.
PN 23-MAR-2002.
XX 12-SEP-2001; 2001WO-GB04094.
PF 12-SEP-2001; 2000GB-0022333.
PR 12-SEP-2000; 2000GB-0022333.
XX (GLAX) GLAXO GROUP LTD.
XX Hiles ID, Ellis C;
XX

```



DR WPI: 2002-351896/38.  
 XX  
 PT Identifying agent that modulates interaction between integrin-linked  
 PT kinase and presenilin-1, useful for treating Alzheimer's disease, and  
 PT identifying agent that modulates protein kinase B or gamma secretase  
 PS activity  
 XX  
 PS Disclosure: Page 43-45; 53pp; English.  
 XX  
 CC The invention relates to identifying an agent modulating interaction  
 CC between integrin-linked kinase (ILK) and presenilin-1 (PS1) comprising  
 CC providing ILK, PS1 or their functional variants, as first and second  
 CC components respectively, contacting the components in the presence of  
 CC a test agent under conditions that in absence of the agent would  
 CC permit the components to interact and determining whether the agent is  
 CC capable of modulating the interaction between the components. Also  
 CC included are the method above where the components are protein kinase  
 CC beta (PKB), gamma secretase or their functional variants, test kits  
 CC for carrying out the methods and an agent identified by the methods.  
 CC The methods are used for identifying an agent that modulates  
 CC interaction between ILK and PS1, or PKB or gamma-secretase. The agent has  
 CC therapeutic applications in treating humans or animals, for treating a  
 CC host suffering from a condition associated with an interaction between  
 CC ILK and PS1, activity of PKB, activity of gamma-secretase, apoptosis,  
 CC cancer and Alzheimer's disease. The agent is also useful for  
 CC manufacturing a medicament for use in treatment of the above mentioned  
 CC conditions. The present sequence represents human presenilin-1, PS-1.  
 CC  
 XX  
 SQ Sequence 467 AA:  
 Query Match 78.4%; Score 366; DB 23; Length 467;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTELPAPLSYFQNAQMSDNHLSNTVRSQNDNREROHNDRRSLGHPPLSGRPOGNSR 60  
 DB 1 MTELPAPLSYFQNAQMSDNHLSNTVRSQNDNREROHNDRRSLGHPPLSGRPOGNSR 60  
 QY 61 QVVEDDEDEDELTLKYGAKHYIMLFVPTLCMVVVAVATIKSVSYTRDGLITPTPE 120  
 DB 61 QVVEDDEDEDELTLKYGAKHYIMLFVPTLCMVVVAVATIKSVSYTRDGLITPTPE 120  
 QY 121 DTEYVGORALHSILNAAMISVIYVMTLLVLYRYRCYKVHAWLIISLLLEFFFSPI 180  
 DB 121 DTEYVGORALHSILNAAMISVIYVMTLLVLYRYRCYKVHAWLIISLLLEFFFSPI 180  
 QY 181 YLGEVFKTYNVAVDYITVALLIMNFGVGMISIHMKGPLRLOQAYLIMISALMALVFIKY 240  
 DB 181 YLGEVFKTYNVAVDYITVALLIMNFGVGMISIHMKGPLRLOQAYLIMISALMALVFIKY 240  
 QY 241 LPEMTAMILINISYVDYVAVLCPRGRLMLETAEORNETLFPALITYSTVWMLVNNAE 300  
 DB 241 LPEMTAMILINISYVDYVAVLCPRGRLMLETAEORNETLFPALITYSTVWMLVNNAE 300  
 QY 301 GDPENQRRYSKSKYNAESTERESQDTVAENDGCFSEEMEAQRSHLCPHRSPEESRAA 360  
 DB 301 GDPENQRRYSKSKYNAESTERESQDTVAENDGCFSEEMEAQRSHLCPHRSPEESRAA 360  
 QY 361 VOELSSSILAGEDPERGVKGLGDFIFYSVLVGKASATASGDMWTTTACFAVAIIIGLCL 420  
 DB 361 VOELSSSILAGEDPERGVKGLGDFIFYSVLVGKASATASGDMWTTTACFAVAIIIGLCL 420  
 QY 421 TLLLLAIFKKALPALPISTTRGLVYFATDYLVOPFMOLAHOFYI 467  
 DB 421 TLLLLAIFKKALPALPISTTRGLVYFATDYLVOPFMOLAHOFYI 467

RESULT 14  
 AAE17051  
 ID AAE17051 standard; Protein; 467 AA.  
 XX  
 AC AAE17051;  
 XX

PT 18-APR-2002 (first entry)  
 XX  
 DE Human mutant presenilin 1 (PS1) wild type protein.  
 XX  
 KW Human; presenilin 1; PS1; amyloid precursor protein; APP; drug screening;  
 KW Alzheimer's disease; Parkinson's disease; multiple sclerosis; stroke;  
 KW Huntington's disease; amyotrophic lateral sclerosis; Picks disease;  
 KW head injury disease; frontal lobe dementia; cerebellar degeneration;  
 KW ischaemic injury; schizophrenia.  
 KW  
 OS Homo sapiens.  
 XX  
 PN ~~W0300202601-A2.~~  
 XX  
 PD 10-JAN-2002.  
 XX  
 XX 29-JUN-2001; 2001WO-US16508.  
 XX  
 XX 30-JUN-2000; 2000US-215345P.  
 XX  
 PA (PHAA ) PHARMACIA & UPJOHN CO.  
 XX  
 PI Carter DB, Tomasselli AG;  
 XX  
 DR WPI: 2002-140082/18.  
 XX  
 PT Novel isolated mutant presenilin 1 and presenilin 2 polypeptides,  
 PT useful for screening of drugs for treating pathologies associated with  
 PT aberrant amyloid precursor protein processing, such as Alzheimer's  
 PT disease  
 XX  
 XX Disclosure: Fig 1; 80pp; English.  
 PS  
 CC The invention relates to mutant presenilin 1 (PS1) and presenilin 2  
 CC (PS2) polypeptides. Presenilin are involved in the processing of amyloid  
 CC precursor protein (APP) from which major amyloidogenic peptides are  
 CC cleaved. Mutant presenilins are useful for identifying agents that  
 CC modulate amyloid beta-peptide (Abeta) derived peptide production. Mutant  
 CC presenilin is also useful as a target for screening drugs useful in the  
 CC treatment of pathologies associated with aberrant amyloid precursor  
 CC protein processing, such as Alzheimer's disease, Parkinson's disease,  
 CC multiple sclerosis, Huntington's disease, amyotrophic lateral sclerosis,  
 CC head injury disease, Picks disease, frontal lobe dementia, cerebellar  
 CC degeneration, stroke, ischemic injury and schizophrenia. A transgenic  
 CC non-human animal is useful for analysing the interaction between APP and  
 CC mutant presenilin-processing protease in vivo, and for screening anti-  
 CC Alzheimer's disease drugs in vivo. A transgenic non-human  
 CC animal is useful for analysing the interaction between APP and mutant  
 CC presenilin-processing protease in vivo, and for screening anti-  
 CC Alzheimer's disease drugs in vivo. The present sequence is human  
 CC PS1 wild type protein.  
 CC  
 XX  
 SQ Sequence 467 AA:  
 Query Match 78.4%; Score 366; DB 23; Length 467;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTELPAPLSYFQNAQMSDNHLSNTVRSQNDNREROHNDRRSLGHPPLSGRPOGNSR 60  
 DB 1 MTELPAPLSYFQNAQMSDNHLSNTVRSQNDNREROHNDRRSLGHPPLSGRPOGNSR 60  
 QY 61 QVVEDDEDEDELTLKYGAKHYIMLFVPTLCMVVVAVATIKSVSYTRDGLITPTPE 120  
 DB 61 QVVEDDEDEDELTLKYGAKHYIMLFVPTLCMVVVAVATIKSVSYTRDGLITPTPE 120  
 QY 121 DTEYVGORALHSILNAAMISVIYVMTLLVLYRYRCYKVHAWLIISLLLEFFFSPI 180  
 DB 121 DTEYVGORALHSILNAAMISVIYVMTLLVLYRYRCYKVHAWLIISLLLEFFFSPI 180  
 QY 181 YLGEVFKTYNVAVDYITVALLIMNFGVGMISIHMKGPLRLOQAYLIMISALMALVFIKY 240  
 DB 181 YLGEVFKTYNVAVDYITVALLIMNFGVGMISIHMKGPLRLOQAYLIMISALMALVFIKY 240

```

OY 241 LPENTAMILAVISYVDLVAVLCPRKPLMLVETAQERNETLLFPALISSTWVWLVNNAE 300
DB 241 LPENTAMILAVISYVDLVAVLCPRKPLMLVETAQERNETLLFPALISSTWVWLVNNAE 300
OY 301 GDPKQRRVSKNSKYNASTERESEDYVAENDDGFSEEMKQSDSLGPHRSTPESRAA 360
DB 301 GDPKQRRVSKNSKYNASTERESEDYVAENDDGFSEEMKQSDSLGPHRSTPESRAA 360
OY 361 VOELSSSILAGEDEPERGVKLGIDFIYSYLVGKASATASGDMNTTITACFVAIILIGLCL 420
DB 361 VOELSSSILAGEDEPERGVKLGIDFIYSYLVGKASATASGDMNTTITACFVAIILIGLCL 420
OY 421 TLLLLAIFKKALPALPISITFGVYFATDYLVOFMDQLAFHOEYI 467
DB 421 TLLLLAIFKKALPALPISITFGVYFATDYLVOFMDQLAFHOEYI 467

RESULT 15
AAW11840
ID AAW11840 standard; Protein; 463 AA.
AC AAW11840;
XX
XX 07-MAY-1997 (first entry)
DE Early onset Alzheimer's disease (EOAD) splice variant polypeptide.
XX
XX Early onset Alzheimer's disease; EOAD; neurodegenerative disease;
XX diagnosis; therapy; inhibitor; antagonist; antibody.
XX
XX Homo sapiens.
XX
XX WO9703086-A1.
XX
XX 30-JUN-1997.
XX
XX 26-JUN-1996; 96WO-US11064.
XX
XX 18-JUL-1995; 95US-0001501.
XX 13-JUL-1995; 95US-0001142.
XX
XX (UYSF-) UNIV SOUTH FLORIDA.
XX
XX Hardy JA;
XX
XX WPI: 1997-118980/11.
XX N-PSDB; AAT59536.
XX
XX Early onset Alzheimer's disease gene - useful for diagnosing a
XX pre-disposition to Alzheimer's disease
XX
XX Disclosure; Fig 2; 4app; English.
XX
XX A 463-amino acid polypeptide (AAW11840) is the product of a full-
XX length cDNA (AAT59536) of an early onset Alzheimer's disease (EOAD)
XX splice variant gene. A 467-amino acid polypeptide (AAW11839) is the
XX product of a full-length cDNA (AAT59535) of the EOAD gene. The 2
XX polypeptides can be produced in transformed host cells and used to
XX raise antibodies, or to identify antagonist/inhibitor cpds. useful
XX in the treatment of Alzheimer's disease, esp. EOAD.
XX
XX Sequence 463 AA;

Query Match 72.2%; Score 337; DB 18; Length 463;
Best Local Similarity 99.8%; Pred. NO. 0;
Matches 437; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 30 NDNEROEHDNRSLGHPPEPLNSGRPOGNSROYVODEDEDELTKYGAHVIMLFPV 89
DB 26 NDNEROEHDNRSLGHPPEPLNSGRPOGNSROYVODEDEDELTKYGAHVIMLFPV 85
C* 90 TLCAVVVAATIKSVSFYTRKGOLITPTEDTETVGORALHSLINAMISVIVMTIL 149

```

```

DB 86 TLCAVVVAATIKSVSFYTRKGOLITPTEDTETVGORALHSLINAMISVIVMTIL 145
OY 150 LVLYKVCYVIAHAWLISLLLPFSFYLGVEFRTYVAADYITVALLINNLGVG 209
DB 146 LVLYKVCYVIAHAWLISLLLPFSFYLGVEFRTYVAADYITVALLINNLGVG 205
OY 210 MISHMKGPRLQOAYLIMISALMAVFIKYLPEMTAWLIAVISYVDLVAVLCPRKPLR 269
DB 206 MISHMKGPRLQOAYLIMISALMAVFIKYLPEMTAWLIAVISYVDLVAVLCPRKPLR 265
OY 270 MLVETAQERNETLLFPALISSTWVWLVNNAEGDEDAQRVSKNSKYNASTERESEDYVA 329
DB 266 MLVETAQERNETLLFPALISSTWVWLVNNAEGDEDAQRVSKNSKYNASTERESEDYVA 325
OY 330 ENDDGGESEEMKQSDSLGPHRSTPESRAVQELSSSILAGEDEPERGVKLGIDFIY 389
DB 326 ENDDGGESEEMKQSDSLGPHRSTPESRAVQELSSSILAGEDEPERGVKLGIDFIY 385
OY 390 SVLVGKASATASGDMNTTITACFVAIILIGLCLTLLLLAIFKKALPALPISITFGVYFAT 449
DB 386 SVLVGKASATASGDMNTTITACFVAIILIGLCLTLLLLAIFKKALPALPISITFGVYFAT 445
OY 450 DYLVQPFMDQLAFHOEYI 467
DB 446 DYLVQPFMDQLAFHOEYI 463

RESULT 16
AAV20854
ID AAV20854 standard; Protein; 463 AA.
AC AAV20854;
XX
XX 22-JUL-1999 (first entry)
DE Human presenilin 1 wild type protein fragment.
XX
XX Human: beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
XX frameshift mutation; age-related disease; neurodegenerative disorder;
XX Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
XX Huntington's disease; multiple sclerosis; alcoholic liver disease;
XX diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
XX ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
XX neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
XX glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
XX bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
XX high mobility group protein-C; neuroendocrine specific protein A.
XX
XX Homo sapiens.
XX
XX WO9845328-A2.
XX
XX 15-OCT-1998.
XX
XX 02-APR-1998; 98WO-IB00705.
XX
XX 10-APR-1997; 97US-0043163.
XX
XX (UYUT-) RIJCKUNIV UTRECHT.
XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
XX (UYRO-) UNIV ROTTERDAM ERASMUS.
XX
XX Burbach JPH, Grosveld FG, Van Leeuwen FW;
XX WPI: 1998-609901/51.
XX N-PSDB; AAX75761.
XX
XX Diagnosing disease by detecting frameshift mutations in RNA or
XX corresponding protein mutations - used to diagnose cancer and
XX neurological diseases, particularly Alzheimer's disease, and also
XX for treatment and prevention with specific ribozymes or wild-type
XX RNA

```

XX Disclosure: Figure 10: 258bp: English.

PS This invention describes a novel method for the diagnosis of a disease  
 CC caused by, or associated with, an RNA molecule that has a frameshift  
 CC mutation. The method is used to diagnose age-related diseases, especially  
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's  
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,  
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II  
 CC and many others listed) or susceptibility to these disorders. The method  
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,  
 CC at an early stage. It is based on the observation that disease may be  
 CC caused by mutations in RNA rather than DNA. The invention describes the  
 CC use of neuronal system RNA molecules, specifically proteins including  
 CC beta-amyloid precursor protein (beta-Ap), the microtubule associated  
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule  
 CC associated protein 2 (Map2), neurofilament-L, neurofilament-M,  
 CC neurofilament-F, presenilin 1, presenilin II, glial fibrillary acidic  
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma  
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group  
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.

XX Sequence 463 AA:

Query Match 72.2%; Score 337; DB 19; Length 463;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 437: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 30 NDNREOEINDRSLGHPPLSNGRPOGNSROVDEDEDEDELTKYGAHVIMLEFV 89  
 DB 26 NDNREOEINDRSLGHPPLSNGRPOGNSROVDEDEDEDELTKYGAHVIMLEFV 85  
 QY 90 TLGMVVVATIKSVSYTRKDGOLITPTEDTEYVGOALHSILNAAMISIVVMTLL 149  
 DB 86 TLGMVVVATIKSVSYTRKDGOLITPTEDTEYVGOALHSILNAAMISIVVMTLL 145  
 QY 150 LVLYKRYCYKIYHAWLIISLLLEFFSFYLGVEFKYNNVADYITVALLIMNGVVG 209  
 DB 146 LVLYKRYCYKIYHAWLIISLLLEFFSFYLGVEFKYNNVADYITVALLIMNGVVG 205  
 QY 210 MISIHKKGLRLQOAYLIMISALMALVFITKYLPEWTAMILAVISYVDLVAVLCPRGPIR 269  
 DB 206 MISIHKKGLRLQOAYLIMISALMALVFITKYLPEWTAMILAVISYVDLVAVLCPRGPIR 265  
 QY 270 MLVETAQERNETLFPALISSTWVWLVNNAEGDPEAQRVSKSKYNAESTERESODTVA 329  
 DB 266 MLVETAQERNETLFPALISSTWVWLVNNAEGDPEAQRVSKSKYNAESTERESODTVA 325  
 QY 330 ENDDGSESEWEAQRDSHLGPHRSTPESRAAVOELSSILAGEDPERGVKLGIDFIFY 389  
 DB 326 ENDDGSESEWEAQRDSHLGPHRSTPESRAAVOELSSILAGEDPERGVKLGIDFIFY 385  
 QY 390 SYLVGKASATASGDMNTTACFAVALIGCLTLLLAIFKKALPALPISITGVLVFPAT 449  
 DB 386 SYLVGKASATASGDMNTTACFAVALIGCLTLLLAIFKKALPALPISITGVLVFPAT 445  
 QY 450 DYLVOPFMDQLAFHOFYI 467  
 DB 446 DYLVOPFMDQLAFHOFYI 463

RESULT 17

AAW11839 standard: Protein: 467 AA.

XX AAW11839:

DT 07-MAY-1997 (first entry)

XX Human early onset Alzheimer's disease (EOAD) polypeptide.

KW Early Onset Alzheimer's disease: EOAD; neurodegenerative disease:  
 diagnosis: therapy; inhibitor; antagonist; antibody.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FT Misc-difference 26..29  
 FT note="unidentified amino acid residues"

XX W09Y03086-A1.

XX 30-JAN-1997.

XX 26-JUN-1996; 96MO-US11064.

XX 18-JUL-1995; 95US-0001501.

XX 13-JUL-1995; 95US-0001142.

XX (UUSF-) UNITV SOUTH FLORIDA.

XX Hardy JA;

XX WPI: 1997-118980/11.

XX N-PSDB: AAT59535.

PT Early onset Alzheimer's disease gene - useful for diagnosing a  
 PT pre-disposition to Alzheimer's disease  
 PS Disclosure: Fig 1: 44pp: English.

XX A 467-amino acid polypeptide (AAW11839) is the product of a  
 CC full-length cDNA (AAT59535) of the early onset Alzheimer's disease  
 CC (EOAD) gene. A 463-amino acid polypeptide (AAW11840) is the  
 CC product of another full-length cDNA (AAT59535) of an EOAD splice  
 CC variant gene. The 2 polypeptides can be produced in transformed  
 CC host cells and used to raise antibodies, or to identify  
 CC antagonist/inhibitor cpds., useful in the treatment of Alzheimer's  
 CC disease, esp. EOAD.

XX Sequence 467 AA:

Query Match 72.2%; Score 337; DB 18; Length 467;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 437: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 30 NDNREOEINDRSLGHPPLSNGRPOGNSROVDEDEDEDELTKYGAHVIMLEFV 89  
 DB 30 NDNREOEINDRSLGHPPLSNGRPOGNSROVDEDEDEDELTKYGAHVIMLEFV 89  
 QY 90 TLGMVVVATIKSVSYTRKDGOLITPTEDTEYVGOALHSILNAAMISIVVMTLL 149  
 DB 90 TLGMVVVATIKSVSYTRKDGOLITPTEDTEYVGOALHSILNAAMISIVVMTLL 149  
 QY 150 LVLYKRYCYKIYHAWLIISLLLEFFSFYLGVEFKYNNVADYITVALLIMNGVVG 209  
 DB 150 LVLYKRYCYKIYHAWLIISLLLEFFSFYLGVEFKYNNVADYITVALLIMNGVVG 209  
 QY 210 MISIHKKGLRLQOAYLIMISALMALVFITKYLPEWTAMILAVISYVDLVAVLCPRGPIR 269  
 DB 210 MISIHKKGLRLQOAYLIMISALMALVFITKYLPEWTAMILAVISYVDLVAVLCPRGPIR 269  
 QY 270 MLVETAQERNETLFPALISSTWVWLVNNAEGDPEAQRVSKSKYNAESTERESODTVA 329  
 DB 270 MLVETAQERNETLFPALISSTWVWLVNNAEGDPEAQRVSKSKYNAESTERESODTVA 329  
 QY 330 ENDDGSESEWEAQRDSHLGPHRSTPESRAAVOELSSILAGEDPERGVKLGIDFIFY 389  
 DB 330 ENDDGSESEWEAQRDSHLGPHRSTPESRAAVOELSSILAGEDPERGVKLGIDFIFY 389  
 QY 390 SYLVGKASATASGDMNTTACFAVALIGCLTLLLAIFKKALPALPISITGVLVFPAT 449  
 DB 390 SYLVGKASATASGDMNTTACFAVALIGCLTLLLAIFKKALPALPISITGVLVFPAT 449  
 QY 450 DYLVOPFMDQLAFHOFYI 467  
 DB 446 DYLVOPFMDQLAFHOFYI 463

DB 450 DYLVQPFMDQLAFHQFYI 467

RESULT 18  
AAW05759 standard; Protein; 429 AA.

ID AAW05759 standard; Protein; 429 AA.

AC AAW05759 standard; Protein; 429 AA.

DE 04-JUN-1998 (first entry)

PS1/429 protein.

Presentin peptide; PS1/429; immunogen; immune response; PS1 gene; Alzheimer's disease; mitochondrial pathology; neurodegeneration; apoptosis.

Homo sapiens.

009746678-A1.

11-DEC-1997.

03-JUN-1997; 97WO-US09272.

18-JUL-1996; 96US-0683315.

06-JUN-1996; 96US-0659296.

(FARB ) BAYER CORP.

Chisholm JC, Davis JN, Drache B.

WPI; 1998-042186/04.

N-PSDB; AAV17357.

DNA encoding presentin peptide PS1/429 and its analogues - useful for diagnosis and treatment of Alzheimer's disease

Claim 24; Fig 1; 77pp; English.

This sequence is the PS1/429 presentin peptide (II) of the invention. Cells transformed with the DNA are used to produce recombinant (II) and analogues, useful e.g. as immunogens for generating an immune response against PS1/429. (II) is a new product of the PS1 gene, mutations in which cause Alzheimer's disease (AD). The nucleic acids are generally useful as probes for detection and quantification of PS1/429, particularly for diagnosis of AD, especially the target sequences that hybridise with probes are isolated for sequencing. Antibodies (Ab) can also be diagnosed at the protein level using Ab as immunoassay reagents. Ab can also be used to identify epitopes and for affinity purification of peptides. Antisense nucleic acid may also be used to regulate expression of the PS1/429 gene, and both nucleic acids and peptides are useful as size markers in electrophoresis, chromatography etc. The transgenic animals are used as models for AD, e.g. for testing drugs. Regulators of the PS1/429 gene or polypeptide can be used to treat e.g. AD or diseases involving mitochondrial pathology, apoptosis and neurodegeneration. Typical regulators are antisense sequences, ribozymes, aptamers, synthetic or natural compounds. (II) may also be used to target other coding sequences to particular cellular locations.

Sequence 429 AA:

Query Match 66.2%; Score 309; DB 19; Length 429;  
Best Local Similarity 99.8%; Pred. No. 1.6e-266;  
Matches 409; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

58 NSROVVEDEDEDELTKYGAHVIMLVPTLCMVVATIKSVSFTYRKDQLIYTP 117  
|||||  
20 NSROVVEDEDEDELTKYGAHVIMLVPTLCMVVATIKSVSFTYRKDQLIYTP 79  
|||||  
118 FTEDTEYVQRAHLSILMAIMISYIVMTLLVLYVKRYKIYHAWLITSSLLLEFF 177  
|||||  
80 FTEDTEYVQRAHLSILMAIMISYIVMTLLVLYVKRYKIYHAWLITSSLLLEFF 139

QY 178 SEIYLGEVEFKTYNVAVDYITVALLIMNLGVGMISIHMKGPLRLQOAYLIMISALMALVF 237  
|||||  
DB 140 SPIYLGVEFKTYNVAVDYITVALLIMNLGVGMISIHMKGPLRLQOAYLIMISALMALVF 199  
|||||  
QY 228 IKYLPETWAMLILAVISYVDLVAVLCPKGPLMLVETAOERNETLFPALIVYSTWMLVN 297  
|||||  
DB 200 IKYLPETWAMLILAVISYVDLVAVLCPKGPLMLVETAOERNETLFPALIVYSTWMLVN 259  
|||||  
QY 298 MAEGDPEAORRYSKSKYNAESTEREODTYAENDGGFSEWEAKRSHLCPHRSPTES 357  
|||||  
DB 260 MAEGDPEAORRYSKSKYNAESTEREODTYAENDGGFSEWEAKRSHLCPHRSPTES 319  
|||||  
QY 358 RAAVDELSSSILAGDEPERGKLGIDFIFYSVLVGRASATASGDMWTITACFVALILG 417  
|||||  
DB 320 RAAVDELSSSILAGDEPERGKLGIDFIFYSVLVGRASATASGDMWTITACFVALILG 379  
|||||  
QY 418 LCLTLLLAIFKKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 467  
|||||  
DB 380 LCLTLLLAIFKKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI 429  
|||||

RESULT 19  
AAW05759 standard; Protein; 467 AA.

ID AAW05759 standard; Protein; 467 AA.

AC AAW05759 standard; Protein; 467 AA.

DE 23-JUL-1997 (first entry)

Presentin-1-1 C410Y mutation.

Presentin-1; human; hPS1-1; hPS1-2; PS-2; integral membrane protein; AD; familial Alzheimer's disease; cerebral haemorrhage; schizophrenia; depression; antibody; gene expression modulator; therapy; muten.

Homo sapiens.

Key Location/Qualifiers  
Modified-site 410 /label=C410Y

WQ634099-A2.

31-QCR-1996.

29-APR-1996; 96WO-CA00263.

31-JUL-1995; 95US-0509359.

28-APR-1995; 95US-0431048.

28-JUN-1995; 95US-0496841.

(HSCR-) HSC RES & DEV LP.

(UTOR ) UNIV TORONTO GOVERNING COUNCIL.

Fraser PE, Rommens JM, St George-Hyslop PH;  
WPI; 1996-497631/49.

New presentin genes - useful for diagnosis, therapy and drug screening of familial Alzheimer's disease, cerebral disorders, etc.

Claim 3; Page -: 178pp; English.

AAW05736-W05760 represent mutated versions of the human presentin-1-1 protein (see AAW05733 for wild type sequence). AAW05734 represents a different wild type form of presentin-1 that results from alternate splicing of the genomic DNA sequence. The presentins are a family of highly conserved integral membrane proteins with a common structural motif, common alternate splicing patterns, and common mutational hot spot regions. Mutations in PS genes are implicated in familial Alzheimer's disease (AD) and possibly other diseases such as cerebral haemorrhage, schizophrenia, depression etc., so detection of mutations in

CC the DNA encoding the wild type sequences can be used for diagnosis of  
 CC these diseases. The wild type proteins, or vectors that express them or  
 CC containing antisense sequences, antibodies selective for these mutant  
 CC forms of the proteins and modulators of PS gene expression are  
 CC potentially useful for treatment of AD etc. Transgenic animals are  
 CC useful as models for drug screening. The antibodies can also be used e.g.  
 CC for affinity purification and in immunoassays.

XX Sequence 467 AA:

Query Match 66.0%; Score 308; DB 17; Length 467;  
 Best Local Similarity 99.8%; Pred. No. 1.7e-295;  
 Matches 408; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MPELPAPLPSYFONAKSENNHLSNTYRSQNDNREHNDKRLSGRPEPLSNRPGNSR 60  
 |||||||  
 Db 1 MPELPAPLPSYFONAKSENNHLSNTYRSQNDNREHNDKRLSGRPEPLSNRPGNSR 60  
 OY 61 QVEODEEEDDELTKYGAHVIMLEFPVTLCAVVVATIKSVFTRKDGOLITPTE 120  
 |||||||  
 Db 61 QVEODEEEDDELTKYGAHVIMLEFPVTLCAVVVATIKSVFTRKDGOLITPTE 120  
 OY 121 DTEVGORALHSLNNAIMISVIVMTLLVLYKRYKVIHAWLIISSLLFFESFI 180  
 |||||||  
 Db 121 DTEVGORALHSLNNAIMISVIVMTLLVLYKRYKVIHAWLIISSLLFFESFI 180  
 OY 181 YLGEVEKTYNVADYITVALLINLGVGMISIHMKGPLRLQOAYLIMSALMALVFIRY 240  
 |||||||  
 Db 181 YLGEVEKTYNVADYITVALLINLGVGMISIHMKGPLRLQOAYLIMSALMALVFIRY 240  
 OY 241 LPEWTAMLLIAVISYDLVAVLCPRKPLRLVETAOERNETLEPALIYSTWVLYNMAE 300  
 |||||||  
 Db 241 LPEWTAMLLIAVISYDLVAVLCPRKPLRLVETAOERNETLEPALIYSTWVLYNMAE 300  
 OY 301 GDEPAORRVSKNSKYNAESTERESQDTVAENDDGFSEMEARQDSHLGHRSTPESRAA 360  
 |||||||  
 Db 301 GDEPAORRVSKNSKYNAESTERESQDTVAENDDGFSEMEARQDSHLGHRSTPESRAA 360  
 OY 361 VOELSSSILAGEDPEERGVKLGDFIFYSVLVGKASATASGMNTTIA 409  
 |||||||  
 Db 361 VOELSSSILAGEDPEERGVKLGDFIFYSVLVGKASATASGMNTTIA 409

RESULT 20

AAW28507  
 ID AAW28507 standard; Protein: 407 AA.

XX AAW28507;

XX 07-DEC-1997 (first entry)

XX Partial AD3 sequence.

XX AD3: AD4/AD3LP; Alzheimer's disease; chromosome; missegregation;  
 KW presentin; inhibitor; AD; trisomy 21.

XX Homo sapiens.

XX Location/Qualifiers

FT key 86 /label= mutation

FT MISC-difference 86 /note= "M -> L"

FT MISC-difference 103 /label= mutation

FT MISC-difference 186 /note= "H -> R"

FT MISC-difference 226 /label= mutation

FT MISC-difference 350 /label= mutation

FT MISC-difference 350 /label= mutation

FT /note= "C -> Y"

XX W09107213-A2.

XX 27-FEB-1997.

XX 15-AUG-1996; 96WO-US1314.

XX 16-AUG-1995; 95US-0002448.

XX (HARD ) HARVARD SOLLEGE.

XX Li J, Potter H;

XX WPI: 1997-165297/15.

XX N-PSDB: AAT87402.

PT Identifying genes which cause chromosome missegregation - useful for  
 PT identifying causes of and treatments for diseases, e.g. Alzheimer's  
 PT disease, cancer and ageing

PS Disclosure: Fig 1: 77pp; English.

CC Identifying genes which cause improper chromosome segregation,  
 CC screening for inhibitors of chromosome missegregation and processes  
 CC caused by genes encoding chromosome missegregation promoters  
 CC was exemplified using Alzheimer's disease. The sequences  
 CC given in Aat87401 to Aat87426 can be used in the above methods.  
 CC The five mutations indicated in the Features table cosegregate  
 CC with early-onset familial Alzheimer's disease. It is predicted  
 CC that these mutations result in increased levels of cells with  
 CC trisomy 21 in carriers of the mutation compared with non-carriers.

XX Sequence 407 AA:

Query Match 65.5%; Score 306; DB 18; Length 407;  
 Best Local Similarity 99.8%; Pred. No. 1.4e-293;  
 Matches 406; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 61 QVEODEEEDDELTKYGAHVIMLEFPVTLCAVVVATIKSVFTRKDGOLITPTE 120  
 |||||||  
 Db 1 QVEODEEEDDELTKYGAHVIMLEFPVTLCAVVVATIKSVFTRKDGOLITPTE 60

OY 121 DTEVGORALHSLNNAIMISVIVMTLLVLYKRYKVIHAWLIISSLLFFESFI 180  
 |||||||  
 Db 61 DTEVGORALHSLNNAIMISVIVMTLLVLYKRYKVIHAWLIISSLLFFESFI 120

OY 181 YLGEVEKTYNVADYITVALLINLGVGMISIHMKGPLRLQOAYLIMSALMALVFIRY 240  
 |||||||  
 Db 121 YLGEVEKTYNVADYITVALLINLGVGMISIHMKGPLRLQOAYLIMSALMALVFIRY 180

OY 241 LPEWTAMLLIAVISYDLVAVLCPRKPLRLVETAOERNETLEPALIYSTWVLYNMAE 300  
 |||||||  
 Db 181 LPEWTAMLLIAVISYDLVAVLCPRKPLRLVETAOERNETLEPALIYSTWVLYNMAE 240

OY 301 GDEPAORRVSKNSKYNAESTERESQDTVAENDDGFSEMEARQDSHLGHRSTPESRAA 360  
 |||||||  
 Db 241 GDEPAORRVSKNSKYNAESTERESQDTVAENDDGFSEMEARQDSHLGHRSTPESRAA 300

OY 361 VOELSSSILAGEDPEERGVKLGDFIFYSVLVGKASATASGMNTTIA 409  
 |||||||  
 Db 361 VOELSSSILAGEDPEERGVKLGDFIFYSVLVGKASATASGMNTTIA 409

XX AD3: AD4/AD3LP; Alzheimer's disease; chromosome; missegregation;  
 KW presentin; inhibitor; AD; trisomy 21.

XX Homo sapiens.

XX Location/Qualifiers

FT key 86 /label= mutation

FT MISC-difference 86 /note= "M -> L"

FT MISC-difference 103 /label= mutation

FT MISC-difference 186 /note= "H -> R"

FT MISC-difference 226 /label= mutation

FT MISC-difference 350 /label= mutation

FT MISC-difference 350 /label= mutation

RESULT 21

AAW05758  
 ID AAW05758 standard; Protein: 467 AA.

XX AAW05758;

```

XX 23-JUL-1997 (first entry)
DT
XX
DE Presenilin-1 L392V mutation.
XX
KW Presenilin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;
KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW depression; antibody; gene expression modulator; therapy; muten.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 392
FT /label= L392V
XX
XX WO9634099-A2.
XX
XX 31-OCT-1996.
XX
XX 29-APR-1996; 96WO-CA00263.
XX
XX 31-JUL-1995; 95US-0509359.
XX 28-APR-1995; 95US-0431048.
XX 28-JUN-1995; 95US-0496841.
XX
PA (HSCR-) HSC RBS & DEV LP.
PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
PI Fraser PE, Rommens JM, St George-Hyslop PH;
XX WPI: 1996-497631/49.
XX
PT New presenilin genes - useful for diagnosis, therapy and drug
PT screening of familial Alzheimer's disease, cerebral disorders, etc.
XX
PS Claim 3; Page -: 178pp; English.
XX
XX AAW05736-W05760 represent mutated versions of the human presenilin-1-1
XX protein (see AAW05734 for wild type sequence). AAW05734 represents a
XX different wild type form of presenilin-1 that results from alternate
XX splicing of the genomic DNA sequence. The presenilins are a family of
XX highly conserved integral membrane proteins with a common structural
XX motif, common alternate splicing patterns, and common mutational hot
XX spot regions. Mutations in PS genes are implicated in familial
XX Alzheimer's disease (AD) and possibly other diseases such as cerebral
XX haemorrhage, schizophrenia, depression etc., so detection of mutations in
XX the DNA encoding the wild type sequences can be used for diagnosis of
XX these diseases. The wild type proteins, or vectors that express them or
XX containing antisense sequences, antibodies selective for these mutant
XX forms of the proteins and modulators of PS gene expression are
XX useful as models for drug screening. The antibodies can also be used e.g.
XX for affinity purification and in immunoassays.
XX
SQ Sequence 467 AA:

```

```

Query Match 62.1%; Score 290; DB 17; Length 467;
Best Local Similarity 99.7%; Pred. No. 1e-277;
Matches 390; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 MTELPAPISTFONQOMSEDNHLSTNVRSONDREROEHNDRLSLGPEPLSGRPOGNSR 60
DB 1 MTELPAPISTFONQOMSEDNHLSTNVRSONDREROEHNDRLSLGPEPLSGRPOGNSR 60
OY 61 OVVEDDEDEDELTKYGAHYIMLFVPTICMVVAVTIKSVSYTRKDGOLITPTPE 120
DB 61 OVVEDDEDEDELTKYGAHYIMLFVPTICMVVAVTIKSVSYTRKDGOLITPTPE 120
OY 121 DTEYGOARLHSLNAAIMISIVVMTILVLYKYRCYKVIHAMIISLLLPFFSFI 180
DB 121 DTEYGOARLHSLNAAIMISIVVMTILVLYKYRCYKVIHAMIISLLLPFFSFI 180
OY 181 YLGEVFKTYNAVDYITVALLIWNLGVMGISIHMKGPIRLQOAYLIMISALMALVFTRY 240
DB 181 YLGEVFKTYNAVDYITVALLIWNLGVMGISIHMKGPIRLQOAYLIMISALMALVFTRY 240

```

```

DB 181 YLGEVFKTYNAVDYITVALLIWNFGVGMGISIHMKGPIRLQOAYLIMISALMALVFTRY 240
OY 241 LPEWTAMLILAVISVYDIAVAIVLCPPRLMLVETTAQERNETLEFPALITYSTWMLVNMAE 300
DB 241 LPEWTAMLILAVISVYDIAVAIVLCPPRLMLVETTAQERNETLEFPALITYSTWMLVNMAE 300
OY 301 GPPEAQRVRSKNSKYNAESTERESODTVAENDDGFSEMEAROSHLGPHRSTPESRAA 360
DB 301 GPPEAQRVRSKNSKYNAESTERESODTVAENDDGFSEMEAROSHLGPHRSTPESRAA 360
OY 361 VOELSSIIAGEDPEERGVKLGIDFIFYSV 391
DB 361 VOELSSIIAGEDPEERGVKLGIDFIFYSV 391

RESULT 22
AAW05736
ID AAW05736 standard; Protein; 467 AA.
XX
XX AAW05736;
XX
XX 23-JUL-1997 (first entry)
XX
DE Presenilin-1 A79X mutation.
XX
KW Presenilin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;
KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW depression; antibody; gene expression modulator; therapy; muten.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 79
FT /label= A79X
FT /note= "X = unspecified amino acid"
XX
XX WO9634099-A2.
XX
XX 31-OCT-1996.
XX
XX 29-APR-1996; 96WO-CA00263.
XX
XX 31-JUL-1995; 95US-0509359.
XX 28-APR-1995; 95US-0431048.
XX 28-JUN-1995; 95US-0496841.
XX
PA (HSCR-) HSC RBS & DEV LP.
PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
PI Fraser PE, Rommens JM, St George-Hyslop PH;
XX WPI: 1996-497631/49.
XX
PT New presenilin genes - useful for diagnosis, therapy and drug
PT screening of familial Alzheimer's disease, cerebral disorders, etc.
XX
PS Claim 3; Page -: 178pp; English.
XX
XX AAW05736-W05760 represent mutated versions of the human presenilin-1-1
XX protein (see AAW05734 for wild type sequence). AAW05734 represents a
XX different wild type form of presenilin-1 that results from alternate
XX splicing of the genomic DNA sequence. The presenilins are a family of
XX highly conserved integral membrane proteins with a common structural
XX motif, common alternate splicing patterns, and common mutational hot
XX spot regions. Mutations in PS genes are implicated in familial
XX Alzheimer's disease (AD) and possibly other diseases such as cerebral
XX haemorrhage, schizophrenia, depression etc., so detection of mutations in
XX the DNA encoding the wild type sequences can be used for diagnosis of
XX these diseases. The wild type proteins, or vectors that express them or
XX containing antisense sequences, antibodies selective for these mutant
XX forms of the proteins and modulators of PS gene expression are
XX potentially useful for treatment of AD etc. Transgenic animals are

```

CC useful as models for drug screening. The antibodies can also be used e.g.  
 CC for affinity purification and in immunoassays.

Sequence 467 AA:

Query Match 61.5%; Score 287; DB 17; Length 467;  
 Best Local Similarity 99.7%; Pred. No. 9.5e-275;  
 Matches 387; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 80 KHYIMFVVTLCMVVAVTIKSVSYTRKDDQLITPTEDETEVGQALHSILNAIM 139  
 DB 80 KHYIMFVVTLCMVVAVTIKSVSYTRKDDQLITPTEDETEVGQALHSILNAIM 139  
 OY 140 ISYIVMTLLVLYKRCYKVIHAWLIISLLLFPSFYIGEYFKYNNVADYITVA 199  
 DB 140 ISYIVMTLLVLYKRCYKVIHAWLIISLLLFPSFYIGEYFKYNNVADYITVA 199  
 OY 200 LLIWNLGVGMISIHMKGPLRLOQAVLIMISALMALVFTRKYPEWTAMLAVISYDLY 259  
 DB 200 LLIWNLGVGMISIHMKGPLRLOQAVLIMISALMALVFTRKYPEWTAMLAVISYDLY 259  
 OY 260 AVLCPRGLRMLVETAOENETLFPALIVSSTWVWLVNNAEGDPEAQRVSKSKNAES 319  
 DB 260 AVLCPRGLRMLVETAOENETLFPALIVSSTWVWLVNNAEGDPEAQRVSKSKNAES 319  
 OY 320 TERESQDTVAENDGGFSEMEQORDSHLGPHRSTPESRAAVOLSSSILAGDPPERGV 379  
 DB 320 TERESQDTVAENDGGFSEMEQORDSHLGPHRSTPESRAAVOLSSSILAGDPPERGV 379  
 OY 380 KLGIDGFIFYSVLVGRASATASGDMNTTACFVAIIIGLCTLLLAIFKKALPALPIST 439  
 DB 380 KLGIDGFIFYSVLVGRASATASGDMNTTACFVAIIIGLCTLLLAIFKKALPALPIST 439  
 OY 440 TEGIVFPAFDYLVOPFMDQAFHQFYI 467  
 DB 440 TEGIVFPAFDYLVOPFMDQAFHQFYI 467

# RESULT 23

AA05737  
 ID AA05737 standard; Protein; 467 AA.

AC AA05737;  
 DT 23-JUL-1997 (first entry)

DE Presentin-1-1 V82L mutation.

KW Presentin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;  
 KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;  
 KW depression; antibody; gene expression modulator; therapy; muten.

OS Homo sapiens.

FT Key Location/Qualifiers  
 FT Modified-site 82  
 FT /label= V82L

PN W09634099-A2.

PD 31-OCT-1996.

PF 29-APR-1996; 96MO-CA00263.

PR 31-JUL-1995; 95US-0509359.

PR 28-APR-1995; 95US-0431048.

PR 28-JUN-1995; 95US-0496841.

PA (HSCR-) HSC RES & DEV LP.

PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.

PI Fraser PE, Rommens JM, St George-Hyslop PH;

DR WPI; 1996-497631/49.

XX New presentin genes - useful for diagnosis, therapy and drug  
 PT screening of familial Alzheimer's disease, cerebral disorders, etc.

PS Claim 3; Page -; 178pp; English.

CC AA05736-W05760 represent mutated versions of the human presentin-1-1  
 CC protein (see AA05733 for wild type sequence). AA05734 represents a  
 CC different wild type form of presentin-1 that results from alternate  
 CC splicing of the genomic DNA sequence. The presentins are a family of  
 CC highly conserved integral membrane proteins with a common structural  
 CC motif, common alternate splicing patterns, and common mutational hot  
 CC spot regions. Mutations in PS genes are implicated in familial  
 CC Alzheimer's disease (AD) and possibly other diseases such as cerebral  
 CC haemorrhage, schizophrenia, depression etc., so detection of mutations in  
 CC the DNA encoding the wild type sequences can be used for diagnosis of  
 CC these diseases. The wild type proteins, or vectors that express them or  
 CC containing antisense sequences, antibodies selective for these mutant  
 CC forms of the proteins and modulators of PS gene expression are  
 CC potentially useful for treatment of AD etc. Transgenic animals are  
 CC useful as models for drug screening. The antibodies can also be used e.g.  
 CC for affinity purification and in immunoassays.

Sequence 467 AA:

Query Match 60.8%; Score 284; DB 17; Length 467;  
 Best Local Similarity 99.7%; Pred. No. 8.8e-272;  
 Matches 384; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 83 IMFVPTLCMVVAVTIKSVSYTRKDDQLITPTEDETEVGQALHSILNAIMISV 142  
 DB 83 IMFVPTLCMVVAVTIKSVSYTRKDDQLITPTEDETEVGQALHSILNAIMISV 142  
 OY 143 IYVMTLLVLYKRCYKVIHAWLIISLLLFPSFYIGEYFKYNNVADYITVALLI 202  
 DB 143 IYVMTLLVLYKRCYKVIHAWLIISLLLFPSFYIGEYFKYNNVADYITVALLI 202  
 OY 203 WNLGVGMISIHMKGPLRLOQAVLIMISALMALVFTRKYPEWTAMLAVISYDLYAVL 262  
 DB 203 WNLGVGMISIHMKGPLRLOQAVLIMISALMALVFTRKYPEWTAMLAVISYDLYAVL 262  
 OY 263 CPKGPLRMLVETAOENETLFPALIVSSTWVWLVNNAEGDPEAQRVSKSKNAESTER 322  
 DB 263 CPKGPLRMLVETAOENETLFPALIVSSTWVWLVNNAEGDPEAQRVSKSKNAESTER 322  
 OY 323 ESQDTVAENDGGFSEMEQORDSHLGPHRSTPESRAAVOLSSSILAGDPPERGV 382  
 DB 323 ESQDTVAENDGGFSEMEQORDSHLGPHRSTPESRAAVOLSSSILAGDPPERGV 382  
 OY 383 LSGDIFYSVLVGRASATASGDMNTTACFVAIIIGLCTLLLAIFKKALPALPISTFG 442  
 DB 383 LSGDIFYSVLVGRASATASGDMNTTACFVAIIIGLCTLLLAIFKKALPALPISTFG 442  
 OY 443 LVFYPATDYLVOPFMDQAFHQFYI 467  
 DB 443 LVFYPATDYLVOPFMDQAFHQFYI 467

# RESULT 24

AAE17046  
 ID AAE17046 standard; Protein; 467 AA.

XX AAE17046.

DT 18-APR-2002 (first entry)

DE Human mutant presentin 1 (ps1) protein #2.

KW Human; presentin 1; PS1; amyloid precursor protein; APP; drug screening;  
 KW Alzheimer's disease; Parkinson's disease; multiple sclerosis; stroke;  
 KW Huntington's disease; amyotrophic lateral sclerosis; Pick's disease;  
 KW head injury disease; frontal lobe dementia; cerebellar degeneration;

```

KW | ischaemic injury; schizophrenia; mutant; mutein.
XX |
OS | Homo sapiens.
XX | Synthetic.
FH | Key Location/Qualifiers
FT | Misc-difference 207..210 /note="Encoded by GTGCTGTG"
FT | Misc-difference 386 /label="Unknown"
FT | /note="Wild type Phe substituted with Xaa; Encoded by NNN"
FT | Misc-difference 387 /label="Unknown"
FT | /note="Wild type Ile substituted with Xaa; Encoded by NNN"
PN | WO200202601-A2.
XX |
PD | 10-JAN-2002.
XX |
PF | 29-JUN-2001; 2001WO-US16508.
XX |
PR | 30-JUN-2000; 2000US-215345P.
XX |
PA | (PHNA ) PHARMACIA & UPJOHN CO.
PI | Carter DB, Tomasselli AG;
XX |
DR | WPT. 2002-140082/18.
DR | N-PSDB; AAD27444.
XX |
PT | Novel isolated mutant presenilin 1 and presenilin 2 polypeptides,
PT | useful for screening of drugs for treating pathologies associated with
PT | aberrant amyloid precursor protein processing, such as Alzheimer's
PT | disease -
XX |
PS | Claim 15; Page 68-70; 80pp; English.
XX |
CC | The invention relates to mutant presenilin 1 (PS1) and presenilin 2
CC | (PS2) polypeptides. Presenilin are involved in the processing of amyloid
CC | precursor protein (APP) from which major amyloidogenic peptides are
CC | cleaved. Mutant presenilins are useful for identifying agents that
CC | modulate amyloid beta-peptide (A-beta) derived peptide production. Mutant
CC | presenilin is also useful as a target for screening drugs useful in the
CC | treatment of pathologies associated with aberrant amyloid precursor
CC | protein processing, such as Alzheimer's disease, Parkinson's disease,
CC | multiple sclerosis, Huntington's disease, amyotrophic lateral sclerosis,
CC | head injury disease, Pick's disease, frontal lobe dementia, cerebellar
CC | degeneration, stroke, ischaemic injury and schizophrenia. A transgenic
CC | non-human animal is useful for analysing the interaction between APP and
CC | mutant presenilin-processing protease in vivo, and for screening anti-
CC | Alzheimer's disease drugs in vivo. A transgenic non-human
CC | animal is useful for analysing the interaction between APP and mutant
CC | presenilin-processing protease in vivo, and for screening anti-
CC | Alzheimer's disease drugs in vivo. The present sequence is human
CC | mutant PS1 protein.
XX |
SQ | Sequence 467 AA;

```

```

Query Match 60.8%; Score 284; DB 23; Length 467;
Best Local Similarity 99.7%; Pred. No. 8.8e-272;
Matches 384; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY | 1 MTELPAPLSTYQNMQMSDNLHLSNVRSONNREKQENDRSLGHPPLSNGRQGSNR 60
DB | 1 MTELPAPLSTYQNMQMSDNLHLSNVRSONNREKQENDRSLGHPPLSNGRQGSNR 60
QY | 61 QVEDDEDEDELTKYGAHVIMLFVPTLCMVVVATIKSVSPYTRKQDOLITTPTE 120
DB | 61 QVEDDEDEDELTKYGAHVIMLFVPTLCMVVVATIKSVSPYTRKQDOLITTPTE 120
QY | 61 QVEDDEDEDELTKYGAHVIMLFVPTLCMVVVATIKSVSPYTRKQDOLITTPTE 120
DB | 61 QVEDDEDEDELTKYGAHVIMLFVPTLCMVVVATIKSVSPYTRKQDOLITTPTE 120
QY | 121 DTEVGGQRALHSLINAAIMISIVVMTLLVLYKRYCKVIHAWLITSSLLLEFFSFI 180
DB | 121 DTEVGGQRALHSLINAAIMISIVVMTLLVLYKRYCKVIHAWLITSSLLLEFFSFI 180

```

```

DB | 121 DTEVGGQRALHSLINAAIMISIVVMTLLVLYKRYCKVIHAWLITSSLLLEFFSFI 180
QY | 181 YGGEVFKTYNNAVVDITVALILWNLGVGMISIHMKGPLRLQOAVLIMISALMAVFYK 240
DB | 181 YGGEVFKTYNNAVVDITVALILWNLGVGMISIHMKGPLRLQOAVLIMISALMAVFYK 240
QY | 241 LPEWTAMTLIAISYVDIAVAVCPKGPLMLVETAOERNETLFPALYSSIPWVWLVNNAE 300
DB | 241 LPEWTAMTLIAISYVDIAVAVCPKGPLMLVETAOERNETLFPALYSSIPWVWLVNNAE 300
QY | 301 GDPFAQRVSKNSKYNAESTEREODTVAENDGFSFEMEQRDSHLGPHRSTPESRAA 360
DB | 301 GDPFAQRVSKNSKYNAESTEREODTVAENDGFSFEMEQRDSHLGPHRSTPESRAA 360
QY | 361 VOELSSSILAGDPPERGYKLGLD 385
DB | 361 VOELSSSILAGDPPERGYKLGLD 385

RESULT 25
AAW05757 ID AAW05757 standard; Protein; 467 AA.
AC AAW05757;
XX 23-JUL-1997 (first entry)
DE Presenilin-1-1 G384A mutation.
XX
KW Presenilin-1; human; hPS1-1; PS-2; integral membrane protein; AD;
KW Familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW depression; antibody; gene expression modulator; therapy; mutein.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 384 /label="G384A"
XX
PN W09634099-A2.
XX
PD 31-OCT-1996.
XX
PF 29-APR-1996; 96WO-CA00263.
XX
PR 31-JUL-1995; 95US-0509359.
PR 28-APR-1995; 95US-0431048.
PR 28-JUN-1995; 95US-0496841.
XX
PA (HSCR-) HSC RES & DEV LP.
PA (UTOR ) UNIV TORONTO GOVERNING COUNCIL.
XX
PI Fraser PE, Rommens JM, St George-Hyslop PH;
DR WPT; 1996-497631/49.
XX
FT New presenilin genes - useful for diagnosis, therapy and drug
FT screening of Familial Alzheimer's disease, cerebral disorders, etc.
XX
PS Claim 3; Page -: 178pp; English.
XX
CC AAW05736-W05760 represent mutated versions of the human presenilin-1-1
CC protein (see AAW05733 for wild type sequence). AAW05734 represents a
CC different wild type form of presenilin-1 that results from alternate
CC splicing of the genomic DNA sequence. The presenilins are a family of
CC highly conserved integral membrane proteins with a common structural
CC motif, common alternate splicing patterns, and common mutational hot
CC spot regions. Mutations in PS genes are implicated in familial
CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
CC the DNA encoding the wild type sequences can be used for diagnosis of
CC these diseases. The wild type proteins, or vectors that express them or

```



CC containing antisense sequences, antibodies selective for these mutant  
CC forms of the proteins and modulators of PS gene expression are  
CC potentially useful for treatment of AD etc. Transgenic animals are  
CC useful as models for drug screening. The antibodies can also be used e.g.  
CC for affinity purification and in immunoassays.

XX Sequence 467 AA:

Query Match 60.4%: Score 282; DB 17; Length 467;  
Best Local Similarity 99.7%: Pred. No. 8.4e-270;  
Matches 382; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTELPAPLSPFONAMQMSDNHLSNTRVSONDRERQEHNDKRSGLGPEPLSGRPOGNSR 60  
DB 1 MTELPAPLSPFONAMQMSDNHLSNTRVSONDRERQEHNDKRSGLGPEPLSGRPOGNSR 60  
QY 61 QVDEDEDEDELTKAKAHVIMLFVVTLCMNVVAVATIKSVSYTRKDGOLITPTPE 120  
DB 61 QVDEDEDEDELTKAKAHVIMLFVVTLCMNVVAVATIKSVSYTRKDGOLITPTPE 120  
QY 121 DTEYVGOALHSILMAAIMISIVVMTLLVLYKRCYKVIHAWLIISLILFFFSFI 180  
DB 121 DTEYVGOALHSILMAAIMISIVVMTLLVLYKRCYKVIHAWLIISLILFFFSFI 180  
QY 181 YLGEVFKTYNAVDYITVALLIMNGVGMISIHMKGPLRQOAVLIMISALMALVFIKY 240  
DB 181 YLGEVFKTYNAVDYITVALLIMNGVGMISIHMKGPLRQOAVLIMISALMALVFIKY 240  
QY 241 LPETAMILAVISYDVAVLCPRGLRMLVETAQERNETLFPALITSTWVLVNMAL 300  
DB 241 LPETAMILAVISYDVAVLCPRGLRMLVETAQERNETLFPALITSTWVLVNMAL 300  
QY 301 GDPKQKRVSKNSKYNASTERESODTVAENDDGGFSEMEQRDSHLGPHRSTPESRA 360  
DB 301 GDPKQKRVSKNSKYNASTERESODTVAENDDGGFSEMEQRDSHLGPHRSTPESRA 360  
QY 361 VOELSSSILAGEDPEERGVKGL 383  
DB 361 VOELSSSILAGEDPEERGVKGL 383

RESULT 26  
AAM05738  
ID AAM05738 standard; Protein; 467 AA.

XX AAM05738:

XX 23-JUL-1997 (first entry)

XX Presentin-1-1 V96F mutation.

XX Presentin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;

KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;  
depression; antibody; gene expression modulator; therapy; muten.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 96 /label= v96f

XX MO9634099-A2.

XX 31-OCT-1996.

XX 29-APR-1996; 96WO-CA00263.

XX 31-JUL-1995; 95US-0509359.

XX 28-APR-1995; 95US-0431048.

XX 28-JUN-1995; 95US-0496841.

PA (HSCR-) HSC RES & DEV LP.  
(UTOR ) UNIV TORONTO GOVERNING COUNCIL.

XX Fraser PE, Rommens JM, St George-Hyslop PH;  
PI  
XX WPI: 1996-497631/49.

XX New presentin genes - useful for diagnosis, therapy and drug  
PT screening of familial Alzheimer's disease, cerebral disorders, etc.

XX Claim 3; Page -: 178pp; English.

CC AAM05736-W05760 represent mutated versions of the human presentin-1-1  
CC protein (see AAM05733 for wild type sequence). AAM05734 represents a  
CC different wild type form of presentin-1 that results from alternate  
CC splicing of the genomic DNA sequence. The presentins are a family of  
CC highly conserved integral membrane proteins with a common structural  
CC motif, common alternate splicing patterns, and common mutational hot  
CC spot regions. Mutations in PS genes are implicated in familial  
CC Alzheimer's disease (AD) and possibly other diseases such as cerebral  
CC haemorrhage, schizophrenia, depression etc., so detection of mutations in  
CC the DNA encoding the wild type sequences can be used for diagnosis of  
CC these diseases. The wild type proteins, or vectors that express them or  
CC containing antisense sequences, antibodies selective for these mutant  
CC forms of the proteins and modulators of PS gene expression are  
CC potentially useful for treatment of AD etc. Transgenic animals are  
CC useful as models for drug screening. The antibodies can also be used e.g.  
CC for affinity purification and in immunoassays.

XX Sequence 467 AA:

Query Match 57.8%: Score 270; DB 17; Length 467;  
Best Local Similarity 99.7%: Pred. No. 6.1e-258;  
Matches 370; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 97 VATIKSVSYTRKDGOLITPTPTEDTEYVGOALHSILMAAIMISIVVMTLLVLYKY 156  
DB 97 VATIKSVSYTRKDGOLITPTPTEDTEYVGOALHSILMAAIMISIVVMTLLVLYKY 156  
QY 157 RCYKVIHAWLIISLILFFFSFIYLGVEFKTYNAVDYITVALLIMNGVGMISIHMK 216  
DB 157 RCYKVIHAWLIISLILFFFSFIYLGVEFKTYNAVDYITVALLIMNGVGMISIHMK 216  
QY 217 GPLRQOAVLIMISALMALVFIKYLPETAMILAVISYDVAVLCPRGLRMLVETAQ 276  
DB 217 GPLRQOAVLIMISALMALVFIKYLPETAMILAVISYDVAVLCPRGLRMLVETAQ 276  
QY 277 ERNETLFPALITSTWVLVNMALVETAGDPEAQRVSKNSKYNASTERESODTVAENDDGGF 336  
DB 277 ERNETLFPALITSTWVLVNMALVETAGDPEAQRVSKNSKYNASTERESODTVAENDDGGF 336  
QY 337 SEEMEQRDSHLGPHRSTPESRAVOELSSSILAGEDPEERGVKGLGPFITYSVLVGKA 396  
DB 337 SEEMEQRDSHLGPHRSTPESRAVOELSSSILAGEDPEERGVKGLGPFITYSVLVGKA 396  
QY 397 SATASGDWMTTACFVAILIGCLTLLLAIFKKALPALPISITGVLVYFATDVLVOPF 456  
DB 397 SATASGDWMTTACFVAILIGCLTLLLAIFKKALPALPISITGVLVYFATDVLVOPF 456  
QY 457 MDQLAFHOFTY 467  
DB 457 MDQLAFHOFTY 467

RESULT 27

AAM05750  
ID AAM05750 standard; Protein; 467 AA.

XX AAM05750:

XX 23-JUL-1997 (first entry)

XX Presentin-1-1 C263R mutation.

XX Presentin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;

KW	familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
XW	depression; antibody; gene expression modulator; therapy; muten.
XX	
OS	Homo sapiens.
XX	
FT	Key Location/Qualifiers
FT	Modified-site 263
FT	/label=C263R
XX	
PN	W09634009-A2.
XX	
PD	.31-OCT-1996.
XX	
PE	29-APR-1996; 96WO-CA00263.
XX	
PR	31-JUL-1995; 95US-0509359.
PR	28-APR-1995; 95US-0431048.
PR	28-JUN-1995; 95US-0496841.
XX	
PA	(HSCR-) HSC RES & DEV LP.
PA	(UTOR ) UNIV TORONTO GOVERNING COUNCIL.
PI	
PI	Fraser PE, Rommens JM, St George-Hyslop PH;
XX	
DR	WPI: 1996-497631/49.
XX	
P7	New presenilin genes - useful for diagnosis, therapy and drug
PT	screening of familial Alzheimer's disease, cerebral disorders, etc.
XX	
PS	Claim 3; Page -: 178pp; English.
XX	
CC	AAM05736-W05760 represent mutated versions of the human presenilin-1-1
CC	protein (see AAM05733 for wild type sequence). AAM05734 represents a
CC	different wild type form of presenilin-1 that results from alternate
CC	splicing of the genomic DNA sequence. The presenilins are a family of
CC	highly conserved integral membrane proteins with a common structural
CC	motif, common alternate splicing patterns, and common mutational hot
CC	spot regions. Mutations in PS genes are implicated in familial
CC	Alzheimer's disease (AD) and possibly other diseases such as cerebral
CC	haemorrhage, schizophrenia, depression etc., so detection of mutations in
CC	the DNA encoding the wild type sequences can be used for diagnosis of
CC	these diseases. The wild type proteins, or vectors that express them or
CC	containing antisense sequences, antibodies selective for these mutant
CC	forms of the proteins and modulators of PS gene expression are
CC	potentially useful for treatment of AD etc. Transgenic animals are
CC	useful as models for drug screening. The antibodies can also be used e.g.
CC	for affinity purification and in immunoassays.
XX	
XX	
SO	Sequence 467 AA:
Query Match	56.7%; Score 265; DB 17; Length 467;
Best Local Similarity	99.6%; Pred. No. 5.3e-253;
Matches 465;	Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Dy	
OY	1 MTELPAPLSTYQNAOMSEDNHLNSTVRSQNDNRRERQHENDRRSLGHPPLSGRPQNSR 60
Dy	
Dy	1 MTELPAPLSTYQNAOMSEDNHLNSTVRSQNDNRRERQHENDRRSLGHPPLSGRPQNSR 60
OY	61 QVWDDEDEDELTLKYAKAHVIMLFVVVTLCMVVVATIKSVSYTRKDGQLITPTE 120
Dy	
Dy	61 QVWDDEDEDELTLKYAKAHVIMLFVVVTLCMVVVATIKSVSYTRKDGQLITPTE 120
OY	121 DTEVGORALSIINAALMISVIYVMTTLLVLVKRYCYKIHAMLISSLLLPFFSFI 180
Dy	
Dy	121 DTEVGORALHSILNAALMISVIYVMTTLLVLVKRYCYKIHAMLISSLLLPFFSFI 180
OY	181 YLGEVPTYNAVVDYITVALLIINMLGVVGMTSIHWKGFLRLQAVLIMISALMALVFITY 240
Dy	
Dy	181 YLGEVPTYNAVVDYITVALLIINMGVVGMTSIHWKGFLRLQAVLIMISALMALVFITY 240
OY	241 LPEPTAMILLIAVISVDYLVAALCPKGPLRMVLVEAQNENLLFPALLIYSTIMWLNVNAE 300
Dy	
Dy	241 LPEPTAMILLIAVISVDYLVAALVRKGPLRMVLVEAQNENLLFPALLIYSTIMWLNVNAE 300

QY	301	GDENQRPVSKNSKYNKSNESPERESQDTPVAENDDGGFSEEMENAROSHLGPHHSTPESRAA	360
Db	301	GDPEQRVRSKNSKYNKSNESPERESQDTPVAENDDGGFSEEMENAROSHLGPHHSTPESRAA	360
QY	361	VOELSSSTILAEDDEBERGVKLGLEDFFEFYVLVGKASATSGDMNTTACFVAIILGLCL	420
Db	361	VOELSSSTILAEDDEBERGVKLGLEDFFEFYVLVGKASATSGDMNTTACFVAIILGLCL	420
QY	421	TLILLATFKKALPALPISITFGVLVFYADLYLVQPEMDLAFHQFYI	467
Db	421	TLILLATFKKALPALPISITFGVLVFYADLYLVQPEMDLAFHQFYI	467
RESULT 28			
ID	AAW05751	AAW05751 standard; Protein; 467 AA.	
XX			
AC	AAW05751;		
XX			
DT	23-JUL-1997	(first entry)	
XX			
DE	Presenilin-1-1 P264L mutation.		
XX			
KW	Presenilin-1; human; hPS1-1; hPS1-2; PS-2; integral membrane protein; AD		
KW	familial Alzheimer's disease; cerebral hemorrhage; schizophrenia;		
KW	depression; antibody; gene expression modulator; therapy; muteln.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	Modified-site	264	
FT		/label= P264L	
XX			
PN	W09634099-A2.		
XX			
PD	31-OCT-1996.		
XX			
PF	29-APR-1996;	96WC-CA00263.	
XX			
XX	31-JUL-1995;	95US-0509359.	
PR	28-APR-1995;	95US-0431048.	
PR	28-JUN-1995;	95US-0496841.	
XX			
PA	(HSCR-) HSC RES & DEV LP.		
PA	(UTOR ) UNIV TORONTO GOVERNING COUNCIL.		
XX			
P1	Fraser PE, Rommens JM, St George-Hyslop PH;		
DR	WPI; 1996-497631/49.		
XX			
PT	New presenilin genes - useful for diagnosis, therapy and drug		
XX	screening of familial Alzheimer's disease, cerebral disorders, etc.		
PS	Claim 3; Page -; 178pp; English.		
XX			
CC	AAW05736-W05760 represent mutated versions of the human presenilin-1-1		
CC	protein (see AAW05733 for wild type sequence). AAW05734 represents a		
CC	different wild type form of presenilin-1 that results from alternate		
CC	splicing of the genomic DNA sequence. The presenilins are a family of		
CC	highly conserved integral membrane proteins with a common structural		
CC	motif, common alternate splicing patterns, and common mutational hot		
CC	spot regions. Mutations in PS genes are implicated in familial		
CC	Alzheimer's disease (AD) and possibly other diseases such as cerebral		
CC	hemorrhage, schizophrenia, depression etc., so detection of mutations in		
CC	the DNA encoding the wild type sequences can be used for diagnosis of		
CC	these diseases. The wild type proteins, or vectors that express them or		
CC	containing antisense sequences, antibodies selective for these mutant		
CC	forms of the proteins and modulators of PS gene expression are		
CC	potentially useful for treatment of AD etc. Transgenic animals are		
CC	useful as models for drug screening. The antibodies can also be used e.g.		
CC	for affinity purification and in immunoassays.		
XX			

SQ Sequence 467 AA:

Query Match 56.7%; Score 265; DB 17; Length 467;  
 Best Local Similarity 99.6%; Pred. No. 5.3e-253;  
 Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MTELPAPLSYFQNAOMSEDNHLSNTVRSQNDNREROEHNDRSLGHPPELSNGRPOGNSR 60  
 DB 1 MTELPAPLSYFQNAOMSEDNHLSNTVRSQNDNREROEHNDRSLGHPPELSNGRPOGNSR 60  
 OY 61 QVVEDDEDEDELTKYGAKHVIMLFVPTLCMVVVAVATIKSVSFYTRKDGQILITPTFE 120  
 DB 61 QVVEDDEDEDELTKYGAKHVIMLFVPTLCMVVVAVATIKSVSFYTRKDGQILITPTFE 120  
 OY 121 DTEYGORALSHILNAAMISYIVVMTILLVLYKRCYKVIHAWLISSLLLEFFFSFI 180  
 DB 121 DTEYGORALSHILNAAMISYIVVMTILLVLYKRCYKVIHAWLISSLLLEFFFSFI 180  
 OY 181 YLGEVFKTYNNAVDTITVALLIMNGVGMISIHMKGPLRLOQAVLIMISALMALVFITY 240  
 DB 181 YLGEVFKTYNNAVDTITVALLIMNGVGMISIHMKGPLRLOQAVLIMISALMALVFITY 240  
 OY 241 LPEWTAMILAVISYVDLVAVLCPRGLMLVETAQERNETLFPALISSTWVWLNNAE 300  
 DB 241 LPEWTAMILAVISYVDLVAVLCPRGLMLVETAQERNETLFPALISSTWVWLNNAE 300  
 OY 301 GDPAQRVRYSKSKYNAESTERESODTYVAENDDGFSEMEQORSHLCPHRSPTESRA 360  
 DB 301 GDPAQRVRYSKSKYNAESTERESODTYVAENDDGFSEMEQORSHLCPHRSPTESRA 360  
 OY 361 VOELSSSILAGEDPERGVKLGIDGFIYSYLVGKASATASGDMWTTTACFAVAILIGLCL 420  
 DB 361 VOELSSSILAGEDPERGVKLGIDGFIYSYLVGKASATASGDMWTTTACFAVAILIGLCL 420  
 OY 421 TLLLLAIFKKALPALPISITTEGLVFFPATDYLVQPFMDLAFHOFYI 467  
 DB 421 TLLLLAIFKKALPALPISITTEGLVFFPATDYLVQPFMDLAFHOFYI 467

RESULT 29

AAW05752 standard; Protein; 467 AA.

AAW05752;

23-JUL-1997 (first entry)

Presentin-1-1 P267S mutation.

XX Presentin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;  
 KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;  
 KW depression; antibody; gene expression modulator; therapy; muten.

OS Homo sapiens.

XX Key Location/Qualifiers

FH Modified-site 267

XX W09634099-A2.

XX 31-OCT-1996.

XX 29-APR-1996; 96WO-CA00263.

XX 31-JUL-1995; 95US-0509359.

XX 28-APR-1995; 95US-0431048.

XX 28-JUN-1995; 95US-0496841.

XX (HSCR-) HSC RES &amp; DEV LP.

XX (UTOR) UNIV TORONTO GOVERNING COUNCIL.

XX Fraser PE, Rommens JM, St George-Hyslop PH;

XX WPI: 1996-497631/49.

DR New presentin genes - useful for diagnosis, therapy and drug  
 XX screening of familial Alzheimer's disease, cerebral disorders, etc.

PS Claim 3; Page -; 178pp; English.

CC AAW05736-W05760 represent mutated versions of the human presentin-1-1  
 CC protein (see AAW05733 for wild type sequence). AAW05734 represents a  
 CC different wild type form of presentin-1 that results from alternate  
 CC splicing of the genomic DNA sequence. The presentins are a family of  
 CC highly conserved integral membrane proteins with a common structural  
 CC motif, common alternate splicing patterns, and common mutational hot  
 CC spot regions. Mutations in PS genes are implicated in familial  
 CC Alzheimer's disease (AD) and possibly other diseases such as cerebral  
 CC haemorrhage, schizophrenia, depression etc., so detection of mutations in  
 CC the DNA encoding the wild type sequences can be used for diagnosis of  
 CC these diseases. The wild type proteins, or vectors that express them or  
 CC containing antisense sequences, antibodies selective for these mutant  
 CC forms of the proteins and modulators of PS gene expression are  
 CC potentially useful for treatment of AD etc. Transgenic animals are  
 CC useful as models for drug screening. The antibodies can also be used e.g.  
 CC for affinity purification and in immunoassays.

SQ Sequence 467 AA:

Query Match 56.7%; Score 265; DB 17; Length 467;  
 Best Local Similarity 99.6%; Pred. No. 5.3e-253;  
 Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MTELPAPLSYFQNAOMSEDNHLSNTVRSQNDNREROEHNDRSLGHPPELSNGRPOGNSR 60  
 DB 1 MTELPAPLSYFQNAOMSEDNHLSNTVRSQNDNREROEHNDRSLGHPPELSNGRPOGNSR 60  
 OY 61 QVVEDDEDEDELTKYGAKHVIMLFVPTLCMVVVAVATIKSVSFYTRKDGQILITPTFE 120  
 DB 61 QVVEDDEDEDELTKYGAKHVIMLFVPTLCMVVVAVATIKSVSFYTRKDGQILITPTFE 120  
 OY 121 DTEYGORALSHILNAAMISYIVVMTILLVLYKRCYKVIHAWLISSLLLEFFFSFI 180  
 DB 121 DTEYGORALSHILNAAMISYIVVMTILLVLYKRCYKVIHAWLISSLLLEFFFSFI 180  
 OY 181 YLGEVFKTYNNAVDTITVALLIMNGVGMISIHMKGPLRLOQAVLIMISALMALVFITY 240  
 DB 181 YLGEVFKTYNNAVDTITVALLIMNGVGMISIHMKGPLRLOQAVLIMISALMALVFITY 240  
 OY 241 LPEWTAMILAVISYVDLVAVLCPRGLMLVETAQERNETLFPALISSTWVWLNNAE 300  
 DB 241 LPEWTAMILAVISYVDLVAVLCPRGLMLVETAQERNETLFPALISSTWVWLNNAE 300  
 OY 301 GDPAQRVRYSKSKYNAESTERESODTYVAENDDGFSEMEQORSHLCPHRSPTESRA 360  
 DB 301 GDPAQRVRYSKSKYNAESTERESODTYVAENDDGFSEMEQORSHLCPHRSPTESRA 360  
 OY 361 VOELSSSILAGEDPERGVKLGIDGFIYSYLVGKASATASGDMWTTTACFAVAILIGLCL 420  
 DB 361 VOELSSSILAGEDPERGVKLGIDGFIYSYLVGKASATASGDMWTTTACFAVAILIGLCL 420  
 OY 421 TLLLLAIFKKALPALPISITTEGLVFFPATDYLVQPFMDLAFHOFYI 467  
 DB 421 TLLLLAIFKKALPALPISITTEGLVFFPATDYLVQPFMDLAFHOFYI 467

RESULT 30

AAW05753 standard; Protein; 467 AA.

AAW05753;

23-JUL-1997 (first entry)

Presentin-1-1 E280A/G mutation.

XX		Presentin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;
KW		familial Alzheimer's disease; cerebral haemorrhage; schizophrenia,
KX		depression; antibody; gene expression modulator; therapy; muten.
OS	Homo sapiens.	
XX		
XX		
FT	Key	Location/Qualifiers
FT	Modified-site	280
FT	/label=	E280X
FT	/note=	"X = Ala, Gly"
PN		
XX		
PD	WO9634099-A2.	
XX		
PD	31-OCT-1996.	
XX		
PF	29-APR-1996;	96WO-CA00263.
XX		
PR	31-JUL-1995;	95US-0509359.
XX		
PR	28-APR-1995;	95US-0431048.
XX		
PR	28-JUN-1995;	95US-0496841.
XX		
PA	(HSCR-) HSC RES & DEV LP.	
XX	(UTOR) UNIV TORONTO GOVERNING COUNCIL.	
PI	Fraser PE, Rommens JM, St George-Hyslop PH;	
DR	WPI: 1996-497631/49.	
XX		
PT	New presentin genes - useful for diagnosis, therapy and drug	
PS	screening of familial Alzheimer's disease, cerebral disorders, etc.	
XX		
PS	Claim 3; Page -: 178pp; English.	
XX		
CC	AAW05736-W05760 represent mutated versions of the human presentin-1-1	
CC	protein (see AAW05733 for wild type sequence). AAW05734 represents a	
CC	different wild type form of presentin-1 that results from alternate	
CC	splicing of the genomic DNA sequence. The presentins are a family of	
CC	highly conserved integral membrane proteins with a common structural	
CC	motif, common alternate splicing patterns, and common mutational hot	
CC	spot regions. Mutations in PS genes are implicated in familial	
CC	Alzheimer's disease (AD) and possibly other diseases such as cerebral	
CC	haemorrhage, schizophrenia, depression etc., so detection of mutations in	
CC	the DNA encoding the wild type sequences can be used for diagnosis of	
CC	these diseases. The wild type proteins, or vectors that express them or	
CC	containing antisense sequences, antibodies selective for these mutant	
CC	forms of the proteins and modulators of PS gene expression are	
CC	potentially useful for treatment of AD etc. Transgenic animals are	
CC	useful as models for drug screening. The antibodies can also be used e.g.	
CC	for affinity purification and in immunoassays.	
XX		
SQ	Sequence	467 AA:
	Query Match	56.7%; Score 265; DB 17; Length 467;
	Best Local Similarity	99.6%; Pred. No. 5, 3e-253;
	Matches 465; Conservative	0; Mismatches 2; Indels 0; Gaps 0
OY	1	MTELPAPISTPCNAOMSEDNHLSTNYRSQNDNREREGHNDRSLGHEPELSNGRPGNSR 60
DB	1	MTELPAPISYFNAOMSEDNHLSTNYRSQNDNREREHDNRSLGHEPELSNGRPGNSR 60
OY	61	QVEODEDEDEDELTLTKYGAKHYIMLFVPTYLCMVVVVATIKSYSFYTRKGOLITYPFE 120
DB	61	QVEODEDEDEDELTLTKYGAKHYIMLFVPTYLCMVVVVAITKSYSTRKDGLITYPFE 120
OY	121	DTEIYGQRALHSILNAAIMISVIYVWTILLVVLKYRCYKVIAHWLTISSLLLEFFSFL 180
DB	121	DTEIYGQRALHSILNAAIMISVIYVWTILLVVLKYRCYKVIAHWLTISSLLLEFFSFL 180
OY	181	YLGEVFKNYNNAVDTITYALLTWNIGVGVMISIHMKGPRLDQAYLIMSALTALVFIFY 240
DB	181	YLGEVFKNYNNAVDTITYALLTWNIGVGVMISIHMKGPRLDQAYLIMSALTALVFIFY 240

QY	241	LPENWAMILLIVISYVDLVAVLCGKGPLRMIVETAOERNETLEPALYSTWMLVNMAE	300
Db	241	LPENWAMILLIVISYVDLVAVLCGKGPLRMIVETAOERNETLEPALYSTWMLVNMAE	300
QY	301	GDPEARVRVSNKSNKYMESTERESQDPIVAERNDGSGFSEEMEAQRDSHLGPHRSTPESRAA	360
Db	301	GDPEARVRVSNKSNKYMESTERESQDPIVAERNDGSGFSEEMEAQRDSHLGPHRSTPESRAA	360
QY	361	VOELSSSTLAGEDEPBERGVKLGCLDETFYISVLVKRASATRSAGDMNTTACFVAAILGICL	420
Db	361	VOELSSSTLAGEDEPBERGVKLGCLDETFYISVLVKRASATRSAGDMNTTACFVAAILGICL	420
QY	421	TLLLIATFKKALPALPISITFGIVFATQTYLVQPEMDOLAFHQFYI	467
Db	421	TLLLIATFKKALPALPISITFGIVFATQTYLVQPEMDOLAFHQFYI	467
RESULT 31			
AAM05754			
XX	AAM05754 standard; Protein: 467 AA.		
AC	AAM05754:		
XX	23-JUL-1997 (first entry)		
DT	Presenilin-1-1 A285V mutation.		
DE	Presenilin-1: human; hps1-1; hps1-2; ps-2; integral membrane protein; AD; familial Alzheimer's disease; cerebral haemorrhage; schizophrenia; depression; antibody; gene expression modulator; therapy; muteln.		
XX	Homo sapiens.		
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	Modified-site	285	
FT		/label= A285V	
XX	W09634099-A2.		
PN	31-OCT-1996.		
XX	31-OCT-1996.		
PD	31-OCT-1996.		
XX	29-APR-1996: 96WC-CA00263.		
PE	31-JUL-1995: 95US-0509359.		
XX	28-APR-1995: 95US-0431048.		
PR	28-JUN-1995: 95US-0496841.		
XX	(HSCR-) HSC RES & DEV LP.		
PA	(UTOR ) UNIV TORONTO GOVERNING COUNCIL.		
XX	Fraser PE, Rommens JM, St George-Hyslop PH.		
PI	WPI: 1996-497631/49.		
XX	New presenilin genes - useful for diagnosis, therapy and drug		
FT	screening of familial Alzheimer's disease, cerebral disorders, etc.		
XX	Claim 3: Page -: 178pp: English.		
PS	AAM05736-W05760 represent mutated versions of the human presenilin-1-1		
CC	protein (see AAM05733 for wild type sequence). AAM05734 represents a		
CC	different wild type form of presenilin-1 that results from alternate		
CC	splicing of the genomic DNA sequence. The presenilins are a family of		
CC	highly conserved integral membrane proteins with a common structural		
CC	motif, common alternate splicing patterns, and common mutational hot		
CC	spot regions. Mutations in PS genes are implicated in familial		
CC	Alzheimer's disease (AD) and possibly other diseases such as cerebral		
CC	haemorrhage, schizophrenia, depression etc., so detection of mutations		
CC	in the DNA encoding the wild type sequences can be used for diagnosis of		
CC	these diseases. The wild type proteins, or vectors that express them or		
CC	containing antisense sequences, antibodies selective for these mutant		
CC	forms of the proteins and modulators of PS gene expression are		
CC	potentially useful for treatment of AD etc. Transgenic animals are		

CC useful as models for drug screening. The antibodies can also be used e.g.  
CC for affinity purification and in immunoassays.

**SQ Sequence 467 AA;**

Query Match

56.78; Score 265; DB 17; Length 467;

Best Local Similarity 99.68; Pred. NO. 5.3e-253;  
Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	MTELPAPLASYFOUAKMSEDNHLNSNTVASQNDNREOCHNDRSITGHEPILSNRPOGNSR	60
Db	1	MTELPAPLASYFOUAKMSEDNHLNSNTVASQNDNREOCHNDRSITGHEPILSNRPOGNSR	60
Qy	61	QVNEBODEEDELTLKYGAKHVIMLEVPVPLCMVVVATIKSVSFYTRKDGOLITPTE	120
Db	61	QVNEBODEEDELTLKYGAKHVIMLEVPVPLCMVVVATIKSVSFYTRKDGOLITPTE	120
Qy	121	DTEYVGRALHSLTMAIMTSVYVMVIMLVLYKYKCYVTHAMLTISLLFFPSFI	180
Db	121	DTEYVGRALHSLTMAIMTSVYVMVIMLVLYKYKCYVTHAMLTISLLFFPSFI	180
Qy	181	YLGEVFETYNVAVDYITVALLINMLGVGMISIHMKGPRLQOAYLIMISALMALVEIKY	240
Db	181	YLGEVFETYNVAVDYITVALLINMFEVGMISIHMKGPRLQOAYLIMISALMALVEIKY	240
Qy	241	LPEPTAMLTIAVISYDVLVAVLCPKPELRMLVETAOERNETLFPALITYSTTMVLVMAE	300
Db	241	LPEPTAMLTIAVISYDVLVAVLCPKPELRMLVETAOERNETLFPALITYSTTMVLVMAE	300
Qy	301	GDPFAORVAKNSKYNAESTERESODTVAANDDGGFESEMEADRDSHLGHRSPPESRAA	360
Db	301	GDPFAORVAKNSKYNAESTERESODTVAANDDGGFESEMEADRDSHLGHRSPPESRAA	360
Qy	361	VOELSSSIILAGEDPEERGVNLGLGDFITFYSVLVKGKASATASGDMNTTIACFVAIILIGLCL	420
Db	361	VOELSSSIILAGEDPEERGVNLGLGDFITFYSVLVKGKASATASGDMNTTIACFVAIILIGLCL	420
Qy	421	TLLLLAIFKKALPALPLSTIFFGLVYFATDYLVOFPFDOLAFHOFTY	467
Db	421	TLLLLAIFKKALPALPLSTIFFGLVYFATDYLVOFPFDOLAFHOFTY	467

RESULT 32  
AAW05755

AC AAW05755;

DT 23-JUL-1997 (first entry)

DE Presenilin-1 L286V mutation.

KW Presentin-1; human; hpsl-1; hpsl-2; PS-2; integral membrane protein; AD  
 KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;  
 KW depression; antibody; gene expression modulator; therapy; mutin.

OS Homo sapiens.

	Key	Location/Qualifiers
FH	Modified-site	286
FT		

PN WO9634099-A2.

PD 31-OCT-1996.

29-APR-1996;

PR 31-JUL-1995;

PR 28-JUN-1995;

PA (HSCR-) HSC RE

PA (UTOR ) UNIV TORONTO GOVERNING COUNCIL.

PI Fraser PE, Rommens JM, St George-Hyslop PH, ...

DR WPI; 1996-497631/49.

PT New presenilin genes - useful for diagnosis, therapy and drug screening of familial Alzheimer's disease, cerebral disorders, etc

PS Claim 3; Page -; 178pp; English.

CC AAN005736-W05760 represent mutated versions of the human presenilin-1-1  
CC protein (see AAN005733 for wild type sequence). AAN005734 represents a  
CC different wild type form of presenilin-1 that results from alternate  
CC splicing of the genomic DNA sequence. The presenilins are a family of  
CC highly conserved integral membrane proteins with a common structural  
CC motif, common alternate splicing patterns, and common mutational hot  
CC spot regions. Mutations in PS genes are implicated in familial  
CC Alzheimer's disease (AD) and possibly other diseases such as cerebral  
CC haemorrhage, schizophrenia, depression etc., so detection of mutations in  
CC the DNA encoding the wild type sequences can be used for diagnosis of  
CC these diseases. The wild type proteins, or vectors that express them or  
CC containing antisense sequences, antibodies selective for these mutant  
CC forms of the proteins and modulators of PS gene expression are  
CC potentially useful for treatment of AD etc. Transgenic animals are  
CC useful as models for drug screening. The antibodies can also be used e.g.  
CC for affinity purification and in immunoassays.

Sequence 467 AA;

Query Match	56.78;	Score 265;	DB 17;	Length 467;
-------------	--------	------------	--------	-------------

Best Local Similarity 99.6%; Pred. No. 5.3e-253;  
Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0

Qy	1	MTPLPALPSTYONMOMSDNHLSTNVRSONNRREROEHNRRLSGHPEPLSNRPOGNSR	60
Db	1	MTPLPALPSTYONMOMSDNHLSTNVRSONNRREROEHNRRLSGHPEPLSNRPOGNSR	60
Qy	61	QVVEODEEDEDLTLKYGAKHYIMLFVPVTLCAWVVVATIKSVSFYRKDGQLYTPFTE	120
Db	61	QVVEODEEDEDLTLKYGAKHYIMLFVPVTLCAWVVVATIKSVSFYRKDGQLYTPFTE	120
Qy	121	DTEYWGQALHSLNAAIMISYIVMTLLVLYKYRCKYTHAMLLISLLFFFSFI	180
Db	121	DTEYWGQALHSLNAAIMISYIVMTLLVLYKYRCKYTHAMLLISLLFFFSFI	180
Qy	181	YLGVEFKRYNNAVDTYVALLIMNLGVVGMISIMHKGPLRLQOAYLIMISALMALVFIKY	240
Db	181	YLGVEFKRYNNAVDTYVALLIMNLGVVGMISIMHKGPLRLQOAYLIMISALMALVFIKY	240
Qy	241	LPENTAMLLAVISYVDLVAVLCPKGPLRMLVETRAOERNETLPFALIYSTWMLVMAE	300
Db	241	LPENTAMLLAVISYVDLVAVLCPKGPLRMLVETRAOERNETLPFALIYSTWMLVMAE	300
Qy	301	GPPEAORRYSKNSKXNAESTERESODTYAENDDGGFSEMDAODSHLGPHRSTPESRAA	360
Db	301	GPPEAORRYSKNSKXNAESTERESODTYAENDDGGFSEMDAODSHLGPHRSTPESRAA	360
Qy	361	VOELSSSILAGEDPERGVKLGIDGFIFYSVLYGKASATASGDMNTTIACFVALITGLCL	420
Db	361	VOELSSSILAGEDPERGVKLGIDGFIFYSVLYGKASATASGDMNTTIACFVALITGLCL	420
Qy	421	TLLLLATFRKALPALPISITFGVLYFYFATDVLVQPFMDQLAFHOPIYI	467
Db	421	TLLLLATFRKALPALPISITFGVLYFYFATDVLVQPFMDQLAFHOPIYI	467

## RESULT 33

ID AAW05739 standard; Protein; 467 AA

AC AAW05739;

```

DT 23-JUL-1997 (first entry)
XX
DE Presenilin-1-1 Y115H mutation.
XX
KW Presenilin-1: human; hps1-1; hps1-2; integral membrane protein; AD;
KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW depression; antibody; gene expression modulator; therapy; muten.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 115
FT /label= Y115H
XX
PN M09634099-A2.
XX
PD 31-OCT-1996.
XX
PF 29-APR-1996; 96WO-CA00263.
XX
PR 31-JUL-1995; 95US-0509359.
PR 28-APR-1995; 95US-0431048.
PR 28-JUN-1995; 95US-0496841.
XX
PA (HSCR-) HSC RES & DEV LP.
PA (UTOR ) UNIV TORONTO GOVERNING COUNCIL.
XX
PI Fraser PE, Rommens JM, St George-Hyslop PH;
XX
DR WPI; 1996-497631/49.
XX
PT New presenilin genes - useful for diagnosis, therapy and drug
PT screening of familial Alzheimer's disease, cerebral disorders, etc.
XX
XX
PS Claim 3; Page -: 178pp; English.
XX
CC AAM05736-W05760 represent mutated versions of the human presenilin-1-1
CC protein (see AAM05733 for wild type sequence). AAM05734 represents a
CC different wild type form of presenilin-1 that results from alternate
CC splicing of the genomic DNA sequence. The presenilins are a family of
CC highly conserved integral membrane proteins with a common structural
CC motif, common alternate splicing patterns, and common mutational hot
CC spot regions. Mutations in PS genes are implicated in familial
CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
CC the DNA encoding the wild type sequences can be used for diagnosis of
CC these diseases. The wild type proteins, or vectors that express them or
CC containing antisense sequences, antibodies selective for these mutant
CC forms of the proteins and modulators of PS gene expression are
CC potentially useful for treatment of AD etc. Transgenic animals are
CC useful as models for drug screening. The antibodies can also be used e.g.
CC for affinity purification and in immunoassays.
XX
XX
SQ Sequence 467 AA:

```

```

Query Match 56.7%; Score 265; DB 17; Length 467;
Best Local Similarity 99.6%; Pred. No. 5.3e-253;
Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 1 MTELPAPLSTYQNNQMSQEDNLSSTVRSQNDNRERHNDNRSGHPEPLSNGRPOGNSR 60
DB 1 MTELPAPLSTYQNNQMSQEDNLSSTVRSQNDNRERHNDNRSGHPEPLSNGRPOGNSR 60
OY 61 QVVEDEDEDEELTKGAKHVMLEFVPTLCWVVVAVATIKSVFYRKQGLLYPFTPE 120
DB 61 QVVEDEDEDEELTKGAKHVMLEFVPTLCWVVVAVATIKSVFYRKQGLLYPFTPE 120
OY 121 DTEVGORALHSIIMAAIMSVIVMTLLVVLKYRCYKVIHAWLITSSLLFFPSFI 180
DB 121 DTEVGORALHSIIMAAIMSVIVMTLLVVLKYRCYKVIHAWLITSSLLFFPSFI 180
OY 181 YLGEVFTYVAVDYITVALLIMNGVVGMIISHWKPEPLRLOQAYLIMISALMAVLFIKY 240
DB 181 YLGEVFTYVAVDYITVALLIMNGVVGMIISHWKPEPLRLOQAYLIMISALMAVLFIKY 240

```

```

DB 181 YLGEVFTYVAVDYITVALLIMNGVVGMIISHWKPEPLRLOQAYLIMISALMAVLFIKY 240
OY 241 LPENTAMILLIIVISYDVAVLCPKGPLRMLVETAOENETLFPALITYSTWMLVNNAE 300
DB 241 LPENTAMILLIIVISYDVAVLCPKGPLRMLVETAOENETLFPALITYSTWMLVNNAE 300
OY 301 GDPEAQRVRSKNSKYNAESTERESQDTVAENDDGGFSEMEQAQDSHLGPHRSTPESRAA 360
DB 301 GDPEAQRVRSKNSKYNAESTERESQDTVAENDDGGFSEMEQAQDSHLGPHRSTPESRAA 360
OY 361 VOELSSSTLAGEDEPERKVGIGDFTFYSLVVGKASTASGDMNTTACVALLIGICTL 420
DB 361 VOELSSSTLAGEDEPERKVGIGDFTFYSLVVGKASTASGDMNTTACVALLIGICTL 420
OY 421 TLLLAIFPKKALPALPISITFGVFPATDYLVOPFMDQLAFHOEYI 467
DB 421 TLLLAIFPKKALPALPISITFGVFPATDYLVOPFMDQLAFHOEYI 467

RESULT 34
AAM05740
ID AAM05740 standard; Protein; 467 AA.
XX
AC AAM05740;
XX
DT 23-JUL-1997 (first entry)
XX
DE Presenilin-1-1 M139T/V mutation.
XX
KW Presenilin-1: human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;
KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW depression; antibody; gene expression modulator; therapy; muten.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 139
FT /label= M139X
FT /note= "X = Thr, Val"
XX
PN M09634099-A2.
XX
PD 31-OCT-1996.
XX
PF 29-APR-1996; 96WO-CA00263.
XX
PR 31-JUL-1995; 95US-0509359.
PR 28-APR-1995; 95US-0431048.
PR 28-JUN-1995; 95US-0496841.
XX
PA (HSCR-) HSC RES & DEV LP.
PA (UTOR ) UNIV TORONTO GOVERNING COUNCIL.
XX
PI Fraser PE, Rommens JM, St George-Hyslop PH;
XX
DR WPI; 1996-497631/49.
XX
PT New presenilin genes - useful for diagnosis, therapy and drug
PT screening of familial Alzheimer's disease, cerebral disorders, etc.
XX
XX
PS Claim 3; Page -: 178pp; English.
XX
CC AAM05736-W05760 represent mutated versions of the human presenilin-1-1
CC protein (see AAM05733 for wild type sequence). AAM05734 represents a
CC different wild type form of presenilin-1 that results from alternate
CC splicing of the genomic DNA sequence. The presenilins are a family of
CC highly conserved integral membrane proteins with a common structural
CC motif, common alternate splicing patterns, and common mutational hot
CC spot regions. Mutations in PS genes are implicated in familial
CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
CC the DNA encoding the wild type sequences can be used for diagnosis of
CC these diseases. The wild type proteins, or vectors that express them or

```

CC containing antisense sequences, antibodies selective for these mutant  
 CC forms of the proteins and modulators of PS gene expression are  
 CC potentially useful for treatment of AD etc. Transgenic animals are  
 CC useful as models for drug screening. The antibodies can also be used e.g.  
 CC for affinity purification and in immunoassays.

XX Sequence 467 AA:

Query Match 56.7%; Score 265; DB 17; Length 467;  
 Best Local Similarity 99.6%; Pred. No. 5.3e-253;  
 Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MTELPAFLSYFQNAQMSNDNHLSTNVRSONDREROEHNDRLSLGHPFLSNGRPOGNSR 60  
 DB 1 MTELPAFLSYFQNAQMSNDNHLSTNVRSONDREROEHNDRLSLGHPFLSNGRPOGNSR 60  
 QY 61 QVEDDEDEDEDELTLYKGAHVIMLFVPTLCMVVVVATIKSVSFYTRKDGILYTPFTE 120  
 DB 61 QVEDDEDEDEDELTLYKGAHVIMLFVPTLCMVVVVATIKSVSFYTRKDGILYTPFTE 120  
 QY 121 DRETGORALHSILNAAIMISYVMTLLVLYKRCYKVIHAWLITISLLLEFFSFI 180  
 DB 121 DRETGORALHSILNAAIXISYVMTLLVLYKRCYKVIHAWLITISLLLEFFSFI 180  
 QY 181 YLGEVFKTYNNAVADYITVALIMNLGVGMISIHMKGPLRLOQAVLIMISALMALVFIKY 240  
 DB 181 YLGEVFKTYNNAVADYITVALIMNLGVGMISIHMKGPLRLOQAVLIMISALMALVFIKY 240  
 QY 241 LPETWAMLLAVISYDLVAVLCPRGRLMLVETAQERNETLFPALITYSTWMLVNMAL 300  
 DB 241 LPETWAMLLAVISYDLVAVLCPRGRLMLVETAQERNETLFPALITYSTWMLVNMAL 300  
 QY 301 GDPEAQRVSKSKYNAESTERESODTVAENDGSEEMEAQRDRLSHLPHRSTPESRAA 360  
 DB 301 GDPEAQRVSKSKYNAESTERESODTVAENDGSEEMEAQRDRLSHLPHRSTPESRAA 360  
 QY 361 VOELSSITLAGEDEPERGVKLGDFIFYSVLVGKASATASGDMWTTTACFAVAILIGLCL 420  
 DB 361 VOELSSITLAGEDEPERGVKLGDFIFYSVLVGKASATASGDMWTTTACFAVAILIGLCL 420  
 QY 421 TLLLLAIFKKALPALPISTFGLVFFATDYLVOPMDLAHFQFYI 467  
 DB 421 TLLLLAIFKKALPALPISTFGLVFFATDYLVOPMDLAHFQFYI 467

# RESULT 35

AAW05741 ID AAW05741 standard; Protein: 467 AA.

XX AAW05741:

XX 23-JUL-1997 (first entry)

XX Presentin-1-1 I143T mutation.

XX Presentin-1: human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;

XX familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;

XX depression; antibody; gene expression modulator; therapy; muten.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Modified-site 143 /label- I143T

XX MO9634099-A2.

XX 31-OCT-1996.

XX 29-APR-1996; 96MO-CA00263.

XX 31-JUL-1995; 95US-0509359.

XX 28-APR-1995; 95US-0431048.

PR 28-JUN-1995; 95US-0496841.

XX (HSCR-) HSC RES & DEV LP.

PA (UTOR ) UNIV TORONTO GOVERNING COUNCIL.

XX Fraser PE, Rommens JM, St George-Hyslop PH;

XX WPI; 1996-497631/49.

XX New presentin genes - useful for diagnosis, therapy and drug

XX screening of familial Alzheimer's disease, cerebral disorders, etc.

XX Claim 3; Page -: 178pp; English.

CC AAW05736-W05760 represent mutated versions of the human presentin-1-1  
 CC protein (see AAW05733 for wild type sequence). AAW05734 represents a  
 CC different wild type form of presentin-1 that results from alternate  
 CC splicing of the genomic DNA sequence. The presentins are a family of  
 CC highly conserved integral membrane proteins with a common structural  
 CC motif, common alternate splicing patterns, and common mutational hot  
 CC spot regions. Mutations in PS genes are implicated in familial  
 CC Alzheimer's disease (AD) and possibly other diseases such as cerebral  
 CC haemorrhage, schizophrenia, depression etc., so detection of mutations in  
 CC the DNA encoding the wild type sequences can be used for diagnosis of  
 CC these diseases. The wild type proteins, or vectors that express them or  
 CC containing antisense sequences, antibodies selective for these mutant  
 CC forms of the proteins and modulators of PS gene expression are  
 CC potentially useful for treatment of AD etc. Transgenic animals are  
 CC useful as models for drug screening. The antibodies can also be used e.g.  
 CC for affinity purification and in immunoassays.

XX Sequence 467 AA:

Query Match 56.7%; Score 265; DB 17; Length 467;  
 Best Local Similarity 99.6%; Pred. No. 5.3e-253;  
 Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MTELPAFLSYFQNAQMSNDNHLSTNVRSONDREROEHNDRLSLGHPFLSNGRPOGNSR 60  
 DB 1 MTELPAFLSYFQNAQMSNDNHLSTNVRSONDREROEHNDRLSLGHPFLSNGRPOGNSR 60  
 QY 61 QVEDDEDEDEDELTLYKGAHVIMLFVPTLCMVVVVATIKSVSFYTRKDGILYTPFTE 120  
 DB 61 QVEDDEDEDEDELTLYKGAHVIMLFVPTLCMVVVVATIKSVSFYTRKDGILYTPFTE 120  
 QY 121 DRETGORALHSILNAAIMISYVMTLLVLYKRCYKVIHAWLITISLLLEFFSFI 180  
 DB 121 DRETGORALHSILNAAIXISYVMTLLVLYKRCYKVIHAWLITISLLLEFFSFI 180  
 QY 181 YLGEVFKTYNNAVADYITVALIMNLGVGMISIHMKGPLRLOQAVLIMISALMALVFIKY 240  
 DB 181 YLGEVFKTYNNAVADYITVALIMNLGVGMISIHMKGPLRLOQAVLIMISALMALVFIKY 240  
 QY 241 LPETWAMLLAVISYDLVAVLCPRGRLMLVETAQERNETLFPALITYSTWMLVNMAL 300  
 DB 241 LPETWAMLLAVISYDLVAVLCPRGRLMLVETAQERNETLFPALITYSTWMLVNMAL 300  
 QY 301 GDPEAQRVSKSKYNAESTERESODTVAENDGSEEMEAQRDRLSHLPHRSTPESRAA 360  
 DB 301 GDPEAQRVSKSKYNAESTERESODTVAENDGSEEMEAQRDRLSHLPHRSTPESRAA 360  
 QY 361 VOELSSITLAGEDEPERGVKLGDFIFYSVLVGKASATASGDMWTTTACFAVAILIGLCL 420  
 DB 361 VOELSSITLAGEDEPERGVKLGDFIFYSVLVGKASATASGDMWTTTACFAVAILIGLCL 420  
 QY 421 TLLLLAIFKKALPALPISTFGLVFFATDYLVOPMDLAHFQFYI 467  
 DB 421 TLLLLAIFKKALPALPISTFGLVFFATDYLVOPMDLAHFQFYI 467

# RESULT 36

AAW05742 ID AAW05742 standard; Protein: 467 AA.



```

XX AC AAM05742;
XX XX 23-JUL-1997 (first entry)
XX XX
XX DE Presenilin-1-1 M146L/V mutation.
XX XX
XX KW Presenilin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;
XX KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
XX KW depression; antibody; gene expression modulator; therapy; muten.
XX OS Homo sapiens.
XX FH
XX FT Key Location/Qualifiers
XX FT Modified-site 146
XX FT /label= M146X
XX FT /note= "X = Leu, Val"
XX
XX PN W09634099-A2.
XX PD
XX PD 31-OCT-1996.
XX PF
XX PF 29-APR-1996; 96WO-CA00263.
XX PR
XX PR 31-JUL-1995; 95US-0509359.
XX PR 28-APR-1995; 95US-0431048.
XX PR 28-JUN-1995; 95US-0496841.
XX PA (HSCR-) HSC RES & DEV LP.
XX PA (UTOR ) UNIV TORONTO GOVERNING COUNCIL.
XX PI Fraser PE, Rommens JM, St George-Hyslop PH;
XX DR WPI; 1996-497631/49.
XX XX
XX PT New presenilin genes - useful for diagnosis, therapy and drug
XX PT screening of familial Alzheimer's disease, cerebral disorders, etc.
XX
XX PS Claim 3; Page -: 178pp; English.
XX
XX CC AAM05736-W05760 represent mutated versions of the human presenilin-1-1
XX CC protein (see AAM05733 for wild type sequence). AAM05734 represents a
XX CC different wild type form of presenilin-1 that results from alternate
XX CC splicing of the genomic DNA sequence. The presenilins are a family of
XX CC highly conserved integral membrane proteins with a common structural
XX CC motif, common alternate splicing patterns, and common mutational hot
XX CC spot regions. Mutations in PS genes are implicated in familial
XX CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
XX CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
XX CC the DNA encoding the wild type sequences can be used for diagnosis of
XX CC these diseases. The wild type proteins, or vectors that express them or
XX CC containing antisense sequences, antibodies selective for these mutant
XX CC forms of the proteins and modulators of PS gene expression are
XX CC potentially useful for treatment of AD etc. Transgenic animals are
XX CC useful as models for drug screening. The antibodies can also be used e.g.
XX CC for affinity purification and in immunoassays.
XX
XX SQ Sequence 467 AA:

```

```

Query Match 56.7%; Score 265; DB 17; Length 467;
Best Local Similarity 99.6%; Pred. NO. 5.3e-253;
Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

DB 121 DTEVGOBALSHILNAAIMISIVVXTLLLVLYKRYCKVYIHAMLISSLLFFFSFI 180
QY 181 YLGEVFKTYNNAVDYITVALLIMNLGVGMISIHMKGPLRLOQAVLIMISALMLVFTRY 240
DB 181 YLGEVFKTYNNAVDYITVALLIMNLGVGMISIHMKGPLRLOQAVLIMISALMLVFTRY 240
QY 241 LPEWTAMILILAVISYVDLVAVLCPKPLRLMVEAQERNETLFPALIVSSTMVLYNNAE 300
DB 241 LPEWTAMILILAVISYVDLVAVLCPKPLRLMVEAQERNETLFPALIVSSTMVLYNNAE 300
QY 301 GDPFAQRVSKNSKYNASTEREODITYAENDDGFSEEMARQDSHLGPHRSTPESRAA 360
DB 301 GDPFAQRVSKNSKYNASTEREODITYAENDDGFSEEMARQDSHLGPHRSTPESRAA 360
QY 361 VOELSSSTIAGEDPERECVKGIGDFIFYSVLYGKASATASGDNNTTACFVALIGLCL 420
DB 361 VOELSSSTIAGEDPERECVKGIGDFIFYSVLYGKASATASGDNNTTACFVALIGLCL 420
QY 421 TLLLLAIFKKALPALPISITFGLVFFPATDYLVOPFMDQLAFHOEYI 467
DB 421 TLLLLAIFKKALPALPISITFGLVFFPATDYLVOPFMDQLAFHOEYI 467

RESULT 37
AAM05743
ID AAM05743 standard; Protein; 467 AA.
XX
XX AC AAM05743;
XX XX
XX DT 23-JUL-1997 (first entry)
XX
XX DE Presenilin-1-1 H163R/Y mutation.
XX
XX KW Presenilin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;
XX KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
XX KW depression; antibody; gene expression modulator; therapy; muten.
XX
XX OS Homo sapiens.
XX FH
XX FT Key Location/Qualifiers
XX FT Modified-site 163
XX FT /label= H163X
XX FT /note= "X = Arg, Tyr"
XX
XX PN W09634099-A2.
XX PD
XX PD 31-OCT-1996.
XX PF
XX PF 29-APR-1996; 96WO-CA00263.
XX PR
XX PR 31-JUL-1995; 95US-0509359.
XX PR 28-APR-1995; 95US-0431048.
XX PR 28-JUN-1995; 95US-0496841.
XX PA (HSCR-) HSC RES & DEV LP.
XX PA (UTOR ) UNIV TORONTO GOVERNING COUNCIL.
XX PI Fraser PE, Rommens JM, St George-Hyslop PH;
XX DR WPI; 1996-497631/49.
XX XX
XX PT New presenilin genes - useful for diagnosis, therapy and drug
XX PT screening of familial Alzheimer's disease, cerebral disorders, etc.
XX
XX PS Claim 3; Page -: 178pp; English.
XX
XX CC AAM05736-W05760 represent mutated versions of the human presenilin-1-1
XX CC protein (see AAM05733 for wild type sequence). AAM05734 represents a
XX CC different wild type form of presenilin-1 that results from alternate
XX CC splicing of the genomic DNA sequence. The presenilins are a family of
XX CC highly conserved integral membrane proteins with a common structural
XX CC motif, common alternate splicing patterns, and common mutational hot
XX CC spot regions. Mutations in PS genes are implicated in familial

```



CC Alzheimer's disease (AD) and possibly other diseases such as cerebral  
 CC haemorrhage, schizophrenia, depression etc., so detection of mutations in  
 CC the DNA encoding the wild type sequences can be used for diagnosis of  
 CC these diseases. The wild type proteins, or vectors that express them or  
 CC containing antisense sequences, antibodies selective for these mutant  
 CC forms of the proteins and modulators of PS gene expression are  
 CC potentially useful for treatment of AD etc. Transgenic animals are  
 CC useful as models for drug screening. The antibodies can also be used e.g.  
 CC for affinity purification and in immunoassays.

XX Sequence 467 AA:

#### Query Match

56.7%: Score 265; DB 17; Length 467;

Best Local Similarity 99.6%: Pred. No. 5.3e-253;

Matches 465: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MTELPAPLSYFQNAQMSQEDNHLSTNVRSONDRERQEHNDRLSLGHPPLSNGRPGNSR 60  
 DB 1 MTELPAPLSYFQNAQMSQEDNHLSTNVRSONDRERQEHNDRLSLGHPPLSNGRPGNSR 60  
 QY 61 QVVEDEDEDEDELTLYGAKHVIMLFVPTLCMVVVVATIKSVFTRKDGQILYPTPE 120  
 DB 61 QVVEDEDEDEDELTLYGAKHVIMLFVPTLCMVVVVATIKSVFTRKDGQILYPTPE 120  
 QY 121 DRETVGORALHSLNNAIMISYIVVMTLLVLYKRCYKVIHAWLIISSLLFFFSFI 180  
 DB 121 DRETVGORALHSLNNAIMISYIVVMTLLVLYKRCYKVIHAWLIISSLLFFFSFI 180  
 QY 181 YLGEVFKTYNNAVDTYTVALLIWNLGVMISIHMKGPLRLQOAYLIMSALMALVFIRY 240  
 DB 181 YLGEVFKTYNNAVDTYTVALLIWNLGVMISIHMKGPLRLQOAYLIMSALMALVFIRY 240  
 QY 241 LPEWTAMLILAVISYDLVAVLCPKGPLRLVETAOERNETLFPALISSTWVWLVNMAE 300  
 DB 241 LPEWTAMLILAVISYDLVAVLCPKGPLRLVETAOERNETLFPALISSTWVWLVNMAE 300  
 QY 301 GDEAQRVSKNSKYNAESTERESQPTVAENDGCFSEMEAROSHLPHRSTPESRAA 360  
 DB 301 GDEAQRVSKNSKYNAESTERESQPTVAENDGCFSEMEAROSHLPHRSTPESRAA 360  
 QY 361 VOELSSIIAGDEPPEERGVKLGIDFIYFYSVLGKASATASGDMNTTICFAVALIGLCL 420  
 DB 361 VOELSSIIAGDEPPEERGVKLGIDFIYFYSVLGKASATASGDMNTTICFAVALIGLCL 420  
 QY 421 TLLLLAIFKKALPALPISITFGLVFVFATDYLVQPMQDLAFHQFYI 467  
 DB 421 TLLLLAIFKKALPALPISITFGLVFVFATDYLVQPMQDLAFHQFYI 467

#### RESULT 38

AAW05744 standard: Protein; 467 AA.

XX AAW05744:

XX 23-JUL-1997 (first entry)

XX Presentin-1 L171P mutation.

XX Presentin-1: human; hps1-1; hps1-2; ps-2; integral membrane protein; AD;  
 KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;  
 KW depression; antibody; gene expression modulator; therapy; mutuin.

OS Homo sapiens.

XX

XX Key Location/Qualifiers

XX Modified-site 171 /label= L171P

XX MO9634099-A2.

XX 31-OCT-1996.

XX

PF 29-APR-1996; 96MO-CA00263.

XX

PR 31-JUL-1995; 95US-0509359.

XX

PR 28-APR-1995; 95US-0431048.

XX

PR 28-JUN-1995; 95US-0496841.

XX

PA (HSCR-) HSC RES & DEV LP.

XX

PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.

XX

PI Fraser PE, Rommens JM, St George-Hyslop PH;

XX

DR WPI; 1996-497631/49.

XX

PT New presentin genes - useful for diagnosis, therapy and drug

XX

PT screening of familial Alzheimer's disease, cerebral disorders, etc.

XX

XX Claim 3; Page -: 178pp; English.

AAW05736-M05760 represent mutated versions of the human presentin-1-1  
 protein (see AAW05733 for wild type sequence). AAW05734 represents a  
 different wild type form of presentin-1 that results from alternate  
 splicing of the genomic DNA sequence. The presentins are a family of  
 highly conserved integral membrane proteins with a common structural  
 motif, common alternate splicing patterns, and common mutational hot  
 spot regions. Mutations in PS genes are implicated in familial  
 Alzheimer's disease (AD) and possibly other diseases such as cerebral  
 haemorrhage, schizophrenia, depression etc., so detection of mutations in  
 CC the DNA encoding the wild type sequences can be used for diagnosis of  
 CC these diseases. The wild type proteins, or vectors that express them or  
 CC containing antisense sequences, antibodies selective for these mutant  
 CC forms of the proteins and modulators of PS gene expression are  
 CC potentially useful for treatment of AD etc. Transgenic animals are  
 CC useful as models for drug screening. The antibodies can also be used e.g.  
 CC for affinity purification and in immunoassays.

XX Sequence 467 AA:

XX Query Match 56.7%: Score 265; DB 17; Length 467;

XX Best Local Similarity 99.6%: Pred. No. 5.3e-253;

XX Matches 465: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MTELPAPLSYFQNAQMSQEDNHLSTNVRSONDRERQEHNDRLSLGHPPLSNGRPGNSR 60  
 DB 1 MTELPAPLSYFQNAQMSQEDNHLSTNVRSONDRERQEHNDRLSLGHPPLSNGRPGNSR 60  
 QY 61 QVVEDEDEDEDELTLYGAKHVIMLFVPTLCMVVVVATIKSVFTRKDGQILYPTPE 120  
 DB 61 QVVEDEDEDEDELTLYGAKHVIMLFVPTLCMVVVVATIKSVFTRKDGQILYPTPE 120  
 QY 121 DRETVGORALHSLNNAIMISYIVVMTLLVLYKRCYKVIHAWLIISSLLFFFSFI 180  
 DB 121 DRETVGORALHSLNNAIMISYIVVMTLLVLYKRCYKVIHAWLIISSLLFFFSFI 180  
 QY 181 YLGEVFKTYNNAVDTYTVALLIWNLGVMISIHMKGPLRLQOAYLIMSALMALVFIRY 240  
 DB 181 YLGEVFKTYNNAVDTYTVALLIWNLGVMISIHMKGPLRLQOAYLIMSALMALVFIRY 240  
 QY 241 LPEWTAMLILAVISYDLVAVLCPKGPLRLVETAOERNETLFPALISSTWVWLVNMAE 300  
 DB 241 LPEWTAMLILAVISYDLVAVLCPKGPLRLVETAOERNETLFPALISSTWVWLVNMAE 300  
 QY 301 GDEAQRVSKNSKYNAESTERESQPTVAENDGCFSEMEAROSHLPHRSTPESRAA 360  
 DB 301 GDEAQRVSKNSKYNAESTERESQPTVAENDGCFSEMEAROSHLPHRSTPESRAA 360  
 QY 361 VOELSSIIAGDEPPEERGVKLGIDFIYFYSVLGKASATASGDMNTTICFAVALIGLCL 420  
 DB 361 VOELSSIIAGDEPPEERGVKLGIDFIYFYSVLGKASATASGDMNTTICFAVALIGLCL 420  
 QY 421 TLLLLAIFKKALPALPISITFGLVFVFATDYLVQPMQDLAFHQFYI 467  
 DB 421 TLLLLAIFKKALPALPISITFGLVFVFATDYLVQPMQDLAFHQFYI 467

```

RESULT 39
AAW05745
ID |AAW05745 standard; Protein; 467 AA.
XX
AC |AAW05745;
XX
DT |23-JUL-1997 (first entry)
XX
DE |Presentin-1-1 G209V mutation.
XX
KW |Presentin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;
KW |familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW |depression; antibody; gene expression modulator; therapy; mutein.
XX
OS |Homo sapiens.
XX
FH |Key
FH |Modified-site 209
FT |Location/Qualifiers
FT |Label= G209V
XX
PN |WO9634099-A2.
XX
PD |31-OCT-1996.
XX
PE |29-APR-1996; 96WO-CA00263.
XX
PR |31-JUL-1995; 95US-0509359.
PR |28-APR-1995; 95US-0431048.
PR |28-JUN-1995; 95US-0496841.
XX
PA |(HSCR-) HSC RES & DEV LP.
PA |(UTOR ) UNIV TORONTO GOVERNING COUNCIL.
XX
PI |Fraser PE, Rommens JM, St George-Hyslop PH;
XX
DR |WPI; 1996-497631/49.
XX
PT |New presentin genes - useful for diagnosis, therapy and drug
PT |screening of familial Alzheimer's disease, cerebral disorders, etc.
XX
PS |Claim 3; Page -: 178pp; English.
XX
CC |AAW05736-W05760 represent mutated versions of the human presentin-1-1
CC |protein (see AAW05733 for wild type sequence). AAW05734 represents a
CC |different wild type form of presentin-1 that results from alternate
CC |splicing of the genomic DNA sequence. The presentins are a family of
CC |highly conserved integral membrane proteins with a common structural
CC |motif, common alternate splicing patterns, and common mutational hot
CC |spot regions. Mutations in PS genes are implicated in familial
CC |Alzheimer's disease (AD) and possibly other diseases such as cerebral
CC |haemorrhage, schizophrenia, depression etc., so detection of mutations in
CC |the DNA encoding the wild type sequences can be used for diagnosis of
CC |these diseases. The wild type proteins, or vectors that express them or
CC |containing antisense sequences, antibodies selective for these mutant
CC |forms of the proteins and modulators of PS gene expression are
CC |potentially useful for treatment of AD etc. Transgenic animals are
CC |useful as models for drug screening. The antibodies can also be used e.g.
CC |for affinity purification and in immunoassays.
XX
SQ |Sequence 467 AA:

Query Match 56.7%; Score 265; DB 17; Length 467;
Best Local Similarity 99.6%; Pred. No. 5.3e-253;
Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

OY |121 DIETVQQRALHSLTNAIMISVIVMTLLVLYKKYRCYKVIHAWLIISLLFFFSFI 180
DB |121 DIETVQQRALHSLTNAIMISVIVMTLLVLYKKYRCYKVIHAWLIISLLFFFSFI 180
OY |181 YLGEVFKTYNAVDTYTVALLIMNLGVGMISIHMKGPRLQOAYLIMISALMAVFIRY 240
DB |181 YLGEVFKTYNAVDTYTVALLIMNLGVGMISIHMKGPRLQOAYLIMISALMAVFIRY 240
OY |241 LPBWTAFLIAYISYVDLAVLCPPKPLRLMVEETAQERNETLFPALISSTWVLVNMME 300
DB |241 LPBWTAFLIAYISYVDLAVLCPPKPLRLMVEETAQERNETLFPALISSTWVLVNMME 300
OY |301 GPPEAQRVSKNSKYNAESTERESODTVAENDDGGFSEMEARDSHLGPHRSTPESRAA 360
DB |301 GPPEAQRVSKNSKYNAESTERESODTVAENDDGGFSEMEARDSHLGPHRSTPESRAA 360
OY |361 VOELSSSILAGDPPEERGVKLGDFIFYSVLGKASATASGDWNTTACFVAIILIGLCL 420
DB |361 VOELSSSILAGDPPEERGVKLGDFIFYSVLGKASATASGDWNTTACFVAIILIGLCL 420
OY |421 TLLLAIFKKALPALPISITFGLVFFAFDYLVOPEMDLAFHOFYI 467
DB |421 TLLLAIFKKALPALPISITFGLVFFAFDYLVOPEMDLAFHOFYI 467

RESULT 40
AAW05746
ID |AAW05746 standard; Protein; 467 AA.
XX
AC |AAW05746;
XX
DT |23-JUL-1997 (first entry)
XX
DE |Presentin-1-1 I211T mutation.
XX
KW |Presentin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;
KW |familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW |depression; antibody; gene expression modulator; therapy; mutein.
XX
OS |Homo sapiens.
XX
FH |Key
FH |Modified-site 211
FT |Location/Qualifiers
FT |Label= I211T
XX
PN |WO9634099-A2.
XX
PD |31-OCT-1996.
XX
PE |29-APR-1996; 96WO-CA00263.
XX
PR |31-JUL-1995; 95US-0509359.
PR |28-APR-1995; 95US-0431048.
PR |28-JUN-1995; 95US-0496841.
XX
PA |(HSCR-) HSC RES & DEV LP.
PA |(UTOR ) UNIV TORONTO GOVERNING COUNCIL.
XX
PI |Fraser PE, Rommens JM, St George-Hyslop PH;
XX
DR |WPI; 1996-497631/49.
XX
PT |New presentin genes - useful for diagnosis, therapy and drug
PT |screening of familial Alzheimer's disease, cerebral disorders, etc.
XX
PS |Claim 3; Page -: 178pp; English.
XX
CC |AAW05736-W05760 represent mutated versions of the human presentin-1-1
CC |protein (see AAW05733 for wild type sequence). AAW05734 represents a
CC |different wild type form of presentin-1 that results from alternate
CC |splicing of the genomic DNA sequence. The presentins are a family of
CC |highly conserved integral membrane proteins with a common structural

```

CC motif, common alternate splicing patterns, and common mutational hot  
 CC spot regions. Mutations in PS genes are implicated in familial  
 CC Alzheimer's disease (AD) and possibly other diseases such as cerebral  
 CC haemorrhage, schizophrenia, depression etc., so detection of mutations in  
 CC the DNA encoding the wild type sequences can be used for diagnosis of  
 CC these diseases. The wild type proteins, or vectors that express them or  
 CC containing antisense sequences, antibodies selective for these mutant  
 CC forms of the proteins and modulators of PS gene expression are  
 CC potentially useful for treatment of AD etc. Transgenic animals are  
 CC useful as models for drug screening. The antibodies can also be used e.g.  
 CC for affinity purification and in immunoassays.

XX Sequence 467 AA:

Query Match 56.7%; Score 265; DB 17; Length 467;  
 Best Local Similarity 99.6%; Pred. No. 5.3e-253;  
 Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MTELPAPLPSYFONAKSEDNHLSNTVRSONDNREOEHNDRSLGHPPELNGRPOGNSR 60  
 DB 1 MTELPAPLPSYFONAKSEDNHLSNTVRSONDNREOEHNDRSLGHPPELNGRPOGNSR 60  
 OY 61 QVEODEEEDDELTKYGAHVIMLFVPTLCMVVVVATIKSVSFYTRKDGQLIYPTFE 120  
 DB 61 QVEODEEEDDELTKYGAHVIMLFVPTLCMVVVVATIKSVSFYTRKDGQLIYPTFE 120  
 OY 121 DFTVGOALHSLTNAIMISVIVMTILLVLYKRYCKVYHAWLIISLLFFFSFI 180  
 DB 121 DFTVGOALHSLTNAIMISVIVMTILLVLYKRYCKVYHAWLIISLLFFFSFI 180  
 OY 181 YLGEVEKTVNAVADYITVALLINLGVGMISIMKGPLRLQOAYLIMISALMAVFIKY 240  
 DB 181 YLGEVEKTVNAVADYITVALLINLGVGMISIMKGPLRLQOAYLIMISALMAVFIKY 240  
 OY 241 LPEMTAMLIAVISYVDLVAVLCPKGPLRLMVLVETAOERNETLTPALISSTMVLVMAE 300  
 DB 241 LPEMTAMLIAVISYVDLVAVLCPKGPLRLMVLVETAOERNETLTPALISSTMVLVMAE 300  
 OY 301 GDEPAORRVSKSKYNAESTERESODTVAENDDGGFSEEMADRDHSLGPHRSTPESRAA 360  
 DB 301 GDEPAORRVSKSKYNAESTERESODTVAENDDGGFSEEMADRDHSLGPHRSTPESRAA 360  
 OY 361 VOELSSSILAGEPDEERGVALGIDFLFYSVLVGKASATASGDMNTTIACFVALILGLCL 420  
 DB 361 VOELSSSILAGEPDEERGVALGIDFLFYSVLVGKASATASGDMNTTIACFVALILGLCL 420  
 OY 421 TLULLAIFKKALPALPISITFGLVFYPADYLVQPFMDQLAFHOQFYI 467  
 DB 421 TLULLAIFKKALPALPISITFGLVFYPADYLVQPFMDQLAFHOQFYI 467

RESULT 41

AAW05747  
 ID AAW05747 standard; Protein; 467 AA.

XX AAW05747;

XX 23-JUL-1997 (first entry)

XX Presentin-1-1 I231T mutation.

XX Presentin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;

KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;

KW depression; antibody; gene expression modulator; therapy; mutin.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 231 /label= I231T

XX MO9634099-A2.

PD 31-OCT-1996.  
 XX  
 PF 29-APR-1996; 96WO-CA00263.  
 XX  
 PR 31-JUL-1995; 95US-0509359.  
 PR 28-APR-1995; 95US-0431048.  
 PR 28-JUN-1995; 95US-0496841.  
 XX  
 PA (HSCR-) HSC RES & DEV LP.  
 PA (UTOR ) UNIT TORONTO GOVERNING COUNCIL.  
 XX  
 PI Fraser PE, Rommens JM, St George-Hyslop PH;  
 XX WPI: 1996-497631/49.  
 PT New presentin genes - useful for diagnosis, therapy and drug  
 PT screening of familial Alzheimer's disease, cerebral disorders, etc.  
 XX  
 PS Claim 3; Page -; 178pp; English.

CC AAW05736-W05760 represent mutated versions of the human presentin-1-1  
 CC protein (see AAW05733 for wild type sequence). AAW05734 represents a  
 CC different wild type form of presentin-1 that results from alternate  
 CC splicing of the genomic DNA sequence. The presentins are a family of  
 CC highly conserved integral membrane proteins with a common structural  
 CC motif, common alternate splicing patterns, and common mutational hot  
 CC spot regions. Mutations in PS genes are implicated in familial  
 CC Alzheimer's disease (AD) and possibly other diseases such as cerebral  
 CC haemorrhage, schizophrenia, depression etc., so detection of mutations in  
 CC the DNA encoding the wild type sequences can be used for diagnosis of  
 CC these diseases. The wild type proteins, or vectors that express them or  
 CC containing antisense sequences, antibodies selective for these mutant  
 CC forms of the proteins and modulators of PS gene expression are  
 CC potentially useful for treatment of AD etc. Transgenic animals are  
 CC useful as models for drug screening. The antibodies can also be used e.g.  
 CC for affinity purification and in immunoassays.

SO Sequence 467 AA:

Query Match 56.7%; Score 265; DB 17; Length 467;  
 Best Local Similarity 99.6%; Pred. No. 5.3e-253;  
 Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MTELPAPLPSYFONAKSEDNHLSNTVRSONDNREOEHNDRSLGHPPELNGRPOGNSR 60  
 DB 1 MTELPAPLPSYFONAKSEDNHLSNTVRSONDNREOEHNDRSLGHPPELNGRPOGNSR 60  
 OY 61 QVEODEEEDDELTKYGAHVIMLFVPTLCMVVVVATIKSVSFYTRKDGQLIYPTFE 120  
 DB 61 QVEODEEEDDELTKYGAHVIMLFVPTLCMVVVVATIKSVSFYTRKDGQLIYPTFE 120  
 OY 121 DFTVGOALHSLTNAIMISVIVMTILLVLYKRYCKVYHAWLIISLLFFFSFI 180  
 DB 121 DFTVGOALHSLTNAIMISVIVMTILLVLYKRYCKVYHAWLIISLLFFFSFI 180  
 OY 181 YLGEVEKTVNAVADYITVALLINLGVGMISIMKGPLRLQOAYLIMISALMAVFIKY 240  
 DB 181 YLGEVEKTVNAVADYITVALLINLGVGMISIMKGPLRLQOAYLIMISALMAVFIKY 240  
 OY 241 LPEMTAMLIAVISYVDLVAVLCPKGPLRLMVLVETAOERNETLTPALISSTMVLVMAE 300  
 DB 241 LPEMTAMLIAVISYVDLVAVLCPKGPLRLMVLVETAOERNETLTPALISSTMVLVMAE 300  
 OY 301 GDEPAORRVSKSKYNAESTERESODTVAENDDGGFSEEMADRDHSLGPHRSTPESRAA 360  
 DB 301 GDEPAORRVSKSKYNAESTERESODTVAENDDGGFSEEMADRDHSLGPHRSTPESRAA 360  
 OY 361 VOELSSSILAGEPDEERGVALGIDFLFYSVLVGKASATASGDMNTTIACFVALILGLCL 420  
 DB 361 VOELSSSILAGEPDEERGVALGIDFLFYSVLVGKASATASGDMNTTIACFVALILGLCL 420  
 OY 421 TLULLAIFKKALPALPISITFGLVFYPADYLVQPFMDQLAFHOQFYI 467  
 DB 421 TLULLAIFKKALPALPISITFGLVFYPADYLVQPFMDQLAFHOQFYI 467

Db 421 TLLLAIEKKALPALPISITFGLVFYFATDYLVPFMDQLAFHQFYI 467

RESULT 42  
AAW05748  
ID |AAW05748 standard; Protein; 467 AA.  
XX |  
AC |AAW05748;  
XX |  
DT 23-JUL-1997 (first entry)  
XX |  
DE Presenilin-1-1 A246E mutation.  
XX |  
KW Presentin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;  
KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;  
KW depression; antibody; gene expression modulator; therapy; mutlein.  
XX |  
OS Homo sapiens.  
XX |  
FH Key Location/Qualifiers  
FT Modified-site 246  
FT /label= A246E  
XX |  
PN W09634099-A2.  
XX |  
PD 31-OCT-1996.  
XX |  
PF 29-APR-1996; 96WO-CA00263.  
XX |  
PR 31-JUL-1995; 95US-0509359.  
PR 28-APR-1995; 95US-0431048.  
PR 28-JUN-1995; 95US-0496841.  
XX |  
PA (HSCR-) HSC RES & DEV LP.  
PA (UTOR ) UNIV TORONTO GOVERNING COUNCIL.  
XX |  
PI Fraser PE, Rommens JM, St George-Hyslop PH;  
XX |  
DR WPI; 1996-497631/49.  
XX |  
PT New presenilin genes - useful for diagnosis, therapy and drug  
PT screening of familial Alzheimer's disease, cerebral disorders, etc.  
XX |  
PS Claim 3; Page -: 178pp; English.  
XX |  
XX |  
CC AAW05736-W05760 represent mutated versions of the human presenilin-1-1  
CC protein (see AAW05733 for wild type sequence). AAW05734 represents a  
CC different wild type form of presenilin-1 that results from alternate  
CC splicing of the genomic DNA sequence. The presenilins are a family of  
CC highly conserved integral membrane proteins with a common structural  
CC motif, common alternate splicing patterns, and common mutational hot  
CC spot regions. Mutations in PS genes are implicated in familial  
CC Alzheimer's disease (AD) and possibly other diseases such as cerebral  
CC haemorrhage, schizophrenia, depression etc., so detection of mutations in  
CC the DNA encoding the wild type sequences can be used for diagnosis of  
CC these diseases. The wild type proteins, or vectors that express them or  
CC containing antisense sequences, antibodies selective for these mutant  
CC forms of the proteins and modulators of PS gene expression are  
CC potentially useful for treatment of AD etc. Transgenic animals are  
CC useful as models for drug screening. The antibodies can also be used e.g.  
CC for affinity purification and in immunoassays.  
XX |  
SQ Sequence 467 AA;

Query Match 56.7%; Score 265; DB 17; Length 467;  
Best Local Similarity 99.6%; Pred. No. 5.3e-253;  
Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MTELPAPLSYTONAQMSDNLSTNSQNDNREROEHNDRLSGHPPLSNGRPOGNSR 60  
DB 1 MTELPAPLSYTONAQMSDNLSTNSQNDNREROEHNDRLSGHPPLSNGRPOGNSR 60  
QY 61 QVEODEEEDDELTKYGAHVIMLFVPTLCMVVVVATIKSVSFYTRKQGLIYTPETE 120

Db 61 QVEODEEEDDELTKYGAHVIMLFVPTLCMVVVVATIKSVSFYTRKQGLIYTPETE 120

QY 121 DPEYGORALHSTLNAINMISVTVMFTLLVLYXKRCKYVHAMLISSLLFFFSFI 180  
Db 121 DPEYGORALHSTLNAINMISVTVMFTLLVLYXKRCKYVHAMLISSLLFFFSFI 180  
QY 181 YLGEVFKTYNVAVDYITVALLIMNLGVGMISIMWKGPRLQOAVYIMISALMALVFIKY 240  
Db 181 YLGEVFKTYNVAVDYITVALLIMNLGVGMISIMWKGPRLQOAVYIMISALMALVFIKY 240  
QY 241 LPEWTAMLLAVISYVDLVAVLCPKGRLPMVETAOERNEFTLPALIYSTVMVLVNMAE 300  
Db 241 LPEWTAMLLAVISYVDLVAVLCPKGRLPMVETAOERNEFTLPALIYSTVMVLVNMAE 300  
QY 301 GPPEAORRVSKSKYNAESTERESODTVANDDGGSFEMEAORSHLGPFRSTPESRAA 360  
Db 301 GPPEAORRVSKSKYNAESTERESODTVANDDGGSFEMEAORSHLGPFRSTPESRAA 360  
QY 361 VOELSSILAGEDPPEERGVKLGDFIFYSVLVGKASATASGDMWTTTACFAVAILIGLCL 420  
Db 361 VOELSSILAGEDPPEERGVKLGDFIFYSVLVGKASATASGDMWTTTACFAVAILIGLCL 420  
QY 421 TLLLAIEKKALPALPISITFGLVFYFATDYLVPFMDQLAFHQFYI 467  
Db 421 TLLLAIEKKALPALPISITFGLVFYFATDYLVPFMDQLAFHQFYI 467

RESULT 43  
AAW05749  
ID |AAW05749 standard; Protein; 467 AA.  
XX |  
AC |AAW05749;  
XX |  
DT 23-JUL-1997 (first entry)  
XX |  
DE Presenilin-1-1 A260V mutation.  
XX |  
KW Presentin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;  
KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;  
KW depression; antibody; gene expression modulator; therapy; mutlein.  
XX |  
OS Homo sapiens.  
XX |  
FH Key Location/Qualifiers  
FT Modified-site 260  
FT /label= A260V  
XX |  
PN W09634099-A2.  
XX |  
PD 31-OCT-1996.  
XX |  
PF 29-APR-1996; 96WO-CA00263.  
XX |  
PR 31-JUL-1995; 95US-0509359.  
PR 28-APR-1995; 95US-0431048.  
PR 28-JUN-1995; 95US-0496841.  
XX |  
PA (HSCR-) HSC RES & DEV LP.  
PA (UTOR ) UNIV TORONTO GOVERNING COUNCIL.  
XX |  
PI Fraser PE, Rommens JM, St George-Hyslop PH;  
XX |  
DR WPI; 1996-497631/49.  
XX |  
PT New presenilin genes - useful for diagnosis, therapy and drug  
PT screening of familial Alzheimer's disease, cerebral disorders, etc.  
XX |  
PS Claim 3; Page -: 178pp; English.  
XX |  
XX |  
CC AAW05736-W05760 represent mutated versions of the human presenilin-1-1  
CC protein (see AAW05733 for wild type sequence). AAW05734 represents a  
CC different wild type form of presenilin-1 that results from alternate

CC splicing of the genomic DNA sequence. The presenilins are a family of  
 CC highly conserved integral membrane proteins with a common structural  
 CC motif, common alternate splicing patterns, and common mutational hot  
 CC spot regions. Mutations in PS genes are implicated in familial  
 CC Alzheimer's disease (AD) and possibly other diseases such as cerebral  
 CC haemorrhage, schizophrenia, depression etc., so detection of mutations in  
 CC the DNA encoding the wild type sequences can be used for diagnosis of  
 CC these diseases. The wild type proteins, or vectors that express them or  
 CC containing antisense sequences, antibodies selective for these mutant  
 CC forms of the proteins and modulators of PS gene expression are  
 CC potentially useful for treatment of AD etc. Transgenic animals are  
 CC useful as models for drug screening. The antibodies can also be used e.g.  
 CC for affinity purification and in immunoassays.

CC Sequence 467 AA:

Query Match 56.7%; Score 265; DB 17; Length 467;  
 Best Local Similarity 99.6%; Pred. No. 5.3e-253;  
 Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MTELPAPLSYFQNAOMSEDNHLSNTVRSONDRERQEHNDRLSLGHPPELSNGRPOGNSR 60  
 DB 1 MTELPAPLSYFQNAOMSEDNHLSNTVRSONDRERQEHNDRLSLGHPPELSNGRPOGNSR 60  
 OY 61 QVEQDEDEDEDELTLYGAKHVIMLFVPTLCMVVAVATIKSVSEYTRKDGOLITPTPE 120  
 DB 61 QVEQDEDEDEDELTLYGAKHVIMLFVPTLCMVVAVATIKSVSEYTRKDGOLITPTPE 120  
 OY 121 DTEYGORALHSILNAAMISYIVVMTLLVLYRYCKYVHAWLISSLLLEFFSFI 180  
 DB 121 DTEYGORALHSILNAAMISYIVVMTLLVLYRYCKYVHAWLISSLLLEFFSFI 180  
 OY 181 YLGEVFKTYNNAVVDITYALLIMNGVGMISIHMKGPLRLQOAVLIMSALMAVFIKY 240  
 DB 181 YLGEVFKTYNNAVVDITYALLIMNGVGMISIHMKGPLRLQOAVLIMSALMAVFIKY 240  
 OY 241 LPEWTAMLLIAVISYDVAVALCPKGPLRLMLETAEORNETLFPALIIYSTVWMLVNAE 300  
 DB 241 LPEWTAMLLIAVISYDVAVALCPKGPLRLMLETAEORNETLFPALIIYSTVWMLVNAE 300  
 OY 301 GDPEAQRVSKSKYNAESTERESODTVAENDGGFSEMEQORDSHLGRHSTPESRAA 360  
 DB 301 GDPEAQRVSKSKYNAESTERESODTVAENDGGFSEMEQORDSHLGRHSTPESRAA 360  
 OY 361 VOELSSSILAGDPPERGKVLGLGDFIFYSVLVGKASATASGDMWTTTACFAAILIGLCL 420  
 DB 361 VOELSSSILAGDPPERGKVLGLGDFIFYSVLVGKASATASGDMWTTTACFAAILIGLCL 420  
 OY 421 TLLLAIFKKALPALPISTTGLVFFATDYLVPFMDLAFHOFYI 467  
 DB 421 TLLLAIFKKALPALPISTTGLVFFATDYLVPFMDLAFHOFYI 467

RESULT 44  
 AAW27176  
 ID AAW27176 standard; Protein: 467 AA.

AC AAW27176:

DT 09-DEC-1997 (first entry)  
 DE Human S182 gene, P51 locus, product related to Alzheimer's disease.  
 KW Mutant; antibody; vaccine; Alzheimer's disease.

OS Homo sapiens.

PN MO9708319-A1.

PD 06-MAR-1997.

PF 03-SEP-1996; 96WO-US14114.

XX

PR 30-AUG-1996; 96US-0706344.  
 PR 31-AUG-1995; 95US-0003054.

PA (GEHO) GEN HOSPITAL CORP.

PI Tanzi RE, Masco W;

DR WPI; 1997-179276/16.  
 N-PSDB; AAT85332.

PT Chromosome 14 early-onset familial Alzheimer's disease gene PS1  
 PT mutants - useful for diagnosing likelihood of developing Alzheimer's  
 PT disease, also anti-sense sequences, antibodies and vaccines to delay  
 PT onset

PS Claim 12; Page 72-73; 99pp; English.

CC The present sequence represents the human S182 gene, PS1 locus, product.  
 CC Mutant PS1 produces a gene product that increases the probability of  
 CC Alzheimer's disease. A nucleic acid sequence able to hybridise to  
 CC sequences coding for a mutant PS1 polypeptide can be used as probes for  
 CC diagnosing an increased likelihood of contracting Alzheimer's disease.  
 CC Antibodies against the mutant polypeptide can also be used for this  
 CC purpose. Vectors containing or expressing a nucleic acid molecule,  
 CC protein or antibody specific for mutant PS1 can be administered to a  
 CC patient to reduce the likelihood, or delay the onset, of Alzheimer's  
 CC disease, e.g. anti-sense RNA expression can be used to decrease  
 CC expression of the PS1 peptide. Transgenic animals expressing the  
 CC Alzheimer's disease protein can be used to test candidate therapeutics  
 CC and to investigate the normal role of PS1. The PS1 peptide may also be  
 CC included in pharmaceutical compositions (vaccines) for Alzheimer's  
 CC disease therapy.

SO Sequence 467 AA:

Query Match 56.7%; Score 265; DB 18; Length 467;  
 Best Local Similarity 99.6%; Pred. No. 5.3e-253;  
 Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MTELPAPLSYFQNAOMSEDNHLSNTVRSONDRERQEHNDRLSLGHPPELSNGRPOGNSR 60  
 DB 1 MTELPAPLSYFQNAOMSEDNHLSNTVRSONDRERQEHNDRLSLGHPPELSNGRPOGNSR 60  
 OY 61 QVEQDEDEDEDELTLYGAKHVIMLFVPTLCMVVAVATIKSVSEYTRKDGOLITPTPE 120  
 DB 61 QVEQDEDEDEDELTLYGAKHVIMLFVPTLCMVVAVATIKSVSEYTRKDGOLITPTPE 120  
 OY 121 DTEYGORALHSILNAAMISYIVVMTLLVLYRYCKYVHAWLISSLLLEFFSFI 180  
 DB 121 DTEYGORALHSILNAAMISYIVVMTLLVLYRYCKYVHAWLISSLLLEFFSFI 180  
 OY 181 YLGEVFKTYNNAVVDITYALLIMNGVGMISIHMKGPLRLQOAVLIMSALMAVFIKY 240  
 DB 181 YLGEVFKTYNNAVVDITYALLIMNGVGMISIHMKGPLRLQOAVLIMSALMAVFIKY 240  
 OY 241 LPEWTAMLLIAVISYDVAVALCPKGPLRLMLETAEORNETLFPALIIYSTVWMLVNAE 300  
 DB 241 LPEWTAMLLIAVISYDVAVALCPKGPLRLMLETAEORNETLFPALIIYSTVWMLVNAE 300  
 OY 301 GDPEAQRVSKSKYNAESTERESODTVAENDGGFSEMEQORDSHLGRHSTPESRAA 360  
 DB 301 GDPEAQRVSKSKYNAESTERESODTVAENDGGFSEMEQORDSHLGRHSTPESRAA 360  
 OY 361 VOELSSSILAGDPPERGKVLGLGDFIFYSVLVGKASATASGDMWTTTACFAAILIGLCL 420  
 DB 361 VOELSSSILAGDPPERGKVLGLGDFIFYSVLVGKASATASGDMWTTTACFAAILIGLCL 420  
 OY 421 TLLLAIFKKALPALPISTTGLVFFATDYLVPFMDLAFHOFYI 467  
 DB 421 TLLLAIFKKALPALPISTTGLVFFATDYLVPFMDLAFHOFYI 467

RESULT 45

```

AAW56770
ID AAW56770 standard; Protein; 467 AA.
XX
AC AAW56770;
XX
DT 13-OCT-1998 (first entry)
XX
DE Homo sapiens PS-1.
XX
KW PS-1; presenilin; presenilin-1; PSP-1; Alzheimer's disease;
KW serine protease; neurodegeneration; predisposition; diagnosis.
XX
OS Homo sapiens.
XX
PN EP828003-A2.
XX
PD 11-MAR-1998.
XX
PF 26-AUG-1997; 97EP-0306501.
XX
PR 13-DEC-1996; 96US-0032875.
PR 06-SEP-1996; 96US-0025436.
PR 25-OCT-1996; 96US-0027873.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
PA (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI Browne MJ, Clinkenbeard HE, Creasy CL, Karren EH;
PI Livi GP, Southan CD;
XX
DR WPI: 1998-161101/15.
DR N-PSDB: AAV29525.
XX
PT Nucleic acids encoding human serum protease protein(s) - used for
PT diagnosing pre-disposition to Alzheimer's disease, etc.
XX
PS Example 1: Page 25-26; 65pp; English.
XX
CC The sequence is that of of presenilin PS-1 which was used
CC in the cloning and isolation of the serine protease PSPI.
XX
XX
SO Sequence 467 AA;

Query Match 56.7%; Score 265; DB 19; Length 467;
Best Local Similarity 99.6%; Pred. No. 5.3e-253;
Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MTELPAPLSTYQNMQMSDNLSTVRSONDRKOEHNDRSLGHPPLSNGRPOGNSR 60
DB 1 MTELPAPLSTYQNMQMSDNLSTVRSONDRKOEHNDRSLGHPPLSNGRPOGNSR 60
OY 61 QVEODEDEDEDELTKYCAKHYIMLEVPVTCMVVVVATIKSVSFYTRKQDLTYTPTE 120
DB 61 QVEODEDEDEDELTKYCAKHYIMLEVPVTCMVVVVATIKSVSFYTRKQDLTYTPTE 120
OY 121 DTEVVGQALSHSLNAAIMISIVVMTLLVLYKRYCYVIMHMLTISLLFFESFI 180
DB 121 DTEVVGQALSHSLNAAIMISIVVMTLLVLYKRYCYVIMHMLTISLLFFESFI 180
OY 121 DTEVVGQALSHSLNAAIMISIVVMTLLVLYKRYCYVIMHMLTISLLFFESFI 180
DB 121 DTEVVGQALSHSLNAAIMISIVVMTLLVLYKRYCYVIMHMLTISLLFFESFI 180
OY 181 YLGEVFKYNNAVDYITVALLIMNLGVGMSTHMKGFLRQQAVALMISLMLVPIKY 240
DB 181 YLGEVFKYNNAVDYITVALLIMNLGVGMSTHMKGFLRQQAVALMISLMLVPIKY 240
OY 181 YLGEVFKYNNAVDYITVALLIMNLGVGMSTHMKGFLRQQAVALMISLMLVPIKY 240
DB 181 YLGEVFKYNNAVDYITVALLIMNLGVGMSTHMKGFLRQQAVALMISLMLVPIKY 240
OY 241 LPEWTAMILLAVISYVDLVAVLCRGPRLRMVETAOERNETLFPALITYSSFMVAVLNMAE 300
DB 241 LPEWTAMILLAVISYVDLVAVLCRGPRLRMVETAOERNETLFPALITYSSFMVAVLNMAE 300
OY 241 LPEWTAMILLAVISYVDLVAVLCRGPRLRMVETAOERNETLFPALITYSSFMVAVLNMAE 300
DB 241 LPEWTAMILLAVISYVDLVAVLCRGPRLRMVETAOERNETLFPALITYSSFMVAVLNMAE 300
OY 301 GDPEAORRVSKNSKYNAESTERESQDTVAENDDGFSEEMEAQRDSHLGPRTPESSRAA 360
DB 301 GDPEAORRVSKNSKYNAESTERESQDTVAENDDGFSEEMEAQRDSHLGPRTPESSRAA 360
OY 361 VOELSSSLAGEDEPBERGVKIGLDFIFYSVLVKASATASGDMNTTIACFVAILIGLCL 420
DB 361 VOELSSSLAGEDEPBERGVKIGLDFIFYSVLVKASATASGDMNTTIACFVAILIGLCL 420

```

```

DB 361 VOELSSSLAGEDEPBERGVKIGLDFIFYSVLVKASATASGDMNTTIACFVAILIGLCL 420
OY 421 TLLLAIFRKALPALPISTFGLVFPADYVOPFMDLAFHOFYI 467
DB 421 TLLLAIFRKALPALPISTFGLVFPADYVOPFMDLAFHOFYI 467

RESULT 46
AAE05563
ID AAE05563 standard; Protein; 467 AA.
XX
AC AAE05563;
XX
DT 24-SEP-2001 (first entry)
XX
DE Human presenilin PSI-FAD mutant M146V.
XX
KW Human; Par-4; presenilin; PSI; neuroprotective; nuclear factor kappa B;
KW NF-kappa B; neuronal degeneration; spinal muscular atrophy; paralysis;
KW peripheral neuropathy; motorneuron disorder; neurodegenerative disorder;
KW Parkinson's disease; Meniere's disease; multiple sclerosis; Bell's palsy;
KW Huntington's chorea; Down's syndrome; amyotrophic lateral sclerosis; ALS;
KW nerve deafness; Alzheimer's disease; epilepsy; mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT MISC-difference 146 /note="Wild type Met substituted with Val"
FT W0200151671-A2.
XX
PD 19-JUL-2001.
XX
PF 08-JAN-2001; 2001WO-US00526.
XX
PR 10-JAN-2000; 2000US-0175200.
PR 04-JAN-2001; 2001US-0754949.
XX
PA (SCIO-) SCIOS INC.
XX
PI McCarthy J, Cordell B;
PI WPI: 2001-451872/48.
XX
PT Identifying inhibitors of neuronal degeneration useful for treating
PT e.g. Alzheimer's disease, by determining the ability of a compound to
PT induce nuclear factor kappa B activation, with the involvement of
PT presenilin or Par-4.
XX
PS Example 2: Page -, 66pp; English.
XX
CC The invention relates to human Par-4 protein, presenilin protein (PS1
CC and PS2) and their corresponding DNA molecules. The invention also
CC relates to a method for identifying inhibitors of neuronal degeneration,
CC comprising cotransfecting eukaryotic host cells expressing presenilin
CC (PS), with a par-4 DNA, and an NF-kappa B dependent reporter construct,
CC exposing the cotransfected cells to a candidate molecule and monitoring
CC the ability of the candidate molecule to induce NF-kappa B activation.
CC Presenilin proteins participates in nuclear factor kappa B (NF-kappa B)
CC signalling and activation. The inhibitors of neuronal degeneration
CC are useful for treating neurodegenerative disorders such as Alzheimer's
CC disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's
CC chorea, Down's syndrome, nerve deafness, Meniere's disease and also for
CC treating peripheral neuropathies, motorneuron disorders such as
CC amyotrophic lateral sclerosis (ALS), Bell's palsy and various conditions
CC involving spinal muscular atrophy and paralysis. The present sequence
CC is human presenilin PSI-FAD (familial Alzheimer's disease) mutant.
CC Note: This sequence is not shown in the specification but is derived from
CC human presenilin (PS1) protein (SEQ ID NO: 4) shown in page 60-61 of the
CC specification (AAE05466).
XX

```

SQ Sequence 467 AA:

Query Match 56.7%; Score 265; DB 22; Length 467;  
 Best Local Similarity 99.6%; Pred. No. 5.3e-253;  
 Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MTELPAPLSTYFONAMSEDNHLSNTYVRSONDRERHNDRLSLGHPPLSLNGRPGNSR 60  
 |||||||  
 DB 1 MTELPAPLSTYFONAMSEDNHLSNTYVRSONDRERHNDRLSLGHPPLSLNGRPGNSR 60  
 QY 61 QVVEDEDEDELTLTKYGAHHYIMLFVPTLCMNVVVAITKVSFTTRKDGOLITPPE 120  
 |||||||  
 DB 61 QVVEDEDEDELTLTKYGAHHYIMLFVPTLCMNVVVAITKVSFTTRKDGOLITPPE 120  
 QY 121 DRETGORALHSILNAAIMISYIVVMTLLVLYKRCYKVHAWLIISSLLLEFFSFI 180  
 |||||||  
 DB 121 DRETGORALHSILNAAIMISYIVVMTLLVLYKRCYKVHAWLIISSLLLEFFSFI 180  
 QY 181 YLGEVFKTYNNAVDTYVALLIMNGVGMISIHMKGPLRLQOAVLIMISALMALVFIKY 240  
 |||||||  
 DB 181 YLGEVFKTYNNAVDTYVALLIMNGVGMISIHMKGPLRLQOAVLIMISALMALVFIKY 240  
 QY 241 LPBWTAMLILAVISYDLAVAVICPKGPLMLVETAOERNETLFPALITSTWMLVNMAE 300  
 |||||||  
 DB 241 LPBWTAMLILAVISYDLAVAVICPKGPLMLVETAOERNETLFPALITSTWMLVNMAE 300  
 QY 301 GDEPQORRVSKSKYNAESTERESODTYVAENDDGFSEMEQORSHLCPHRSTPESRA 360  
 |||||||  
 DB 301 GDEPQORRVSKSKYNAESTERESODTYVAENDDGFSEMEQORSHLCPHRSTPESRA 360  
 QY 361 VOELSSSILAGEDPERGVKLGDFIFYSVYVGKASATASGDWNTTACFAAILIGLCL 420  
 |||||||  
 DB 361 VOELSSSILAGEDPERGVKLGDFIFYSVYVGKASATASGDWNTTACFAAILIGLCL 420  
 QY 421 TLLLLAIFKKALPALPISTTGLVFFATDYLVOPMDLAFHOFYI 467  
 |||||||  
 DB 421 TLLLLAIFKKALPALPISTTGLVFFATDYLVOPMDLAFHOFYI 467

RESULT 47

AAE05564 ID AAE05564 standard; Protein: 467 AA.

XX AAE05564;

XX 24-SEP-2001 (first entry)

XX Human presenilin PS1-FAD mutant E280G.

XX Human; Par-4; presenilin; PS1; neuroprotective; nuclear factor kappa B;

XX NF-kappa B; neuronal degeneration; spinal muscular atrophy; paralysis;

XX peripheral neuropathy; motorneuron disorder; neurodegenerative disorder;

XX Parkinson's disease; Meniere's disease; multiple sclerosis; Bell's palsy;

XX Huntington's chorea; Down's syndrome; amyotrophic lateral sclerosis; ALS;

XX nerve deafness; Alzheimer's disease; epilepsy; mutant; mutain.

XX Homo sapiens.

XX Synthetic.

XX Key Location/Qualifiers

XX MISC-difference 280 /note= "Wild type Glu substituted with Gly"

XX WO200151671-A2.

XX 19-JUL-2001.

XX 08-JAN-2001; 2001WO-US00526.

XX 10-JAN-2000; 2000US-0175200.

XX 04-JAN-2001; 2001US-0754949.

XX (SCIO-) SCIOS INC.

XX Mccarthy J, Cordell B;  
 XX WPI: 2001-451872/48.

PT Identifying inhibitors of neuronal degeneration useful for treating  
 PT e.g. Alzheimer's disease, by determining the ability of a compound to  
 PT induce nuclear factor kappa B activation, with the involvement of  
 PT presenilin or Par-4 -

XX Example 2; Page -: 66pp; English.

CC The invention relates to human Par-4 protein, presenilin protein (PS1  
 CC and PS2) and their corresponding DNA molecules. The invention also  
 CC relates to a method for identifying inhibitors of neuronal degeneration,  
 CC comprising cotransfecting eukaryotic host cells expressing presenilin  
 CC (PS), with a Par-4 DNA, and an NF-kappa B dependent reporter construct,  
 CC exposing the cotransfected cells to a candidate molecule and monitoring  
 CC the ability of the candidate molecule to induce NF-kappa B activation.  
 CC Presenilin proteins participate in nuclear factor kappa B (NF-kappa B)  
 CC signalling and activation. The inhibitors of neuronal degeneration  
 CC are useful for treating neurodegenerative disorders such as Alzheimer's  
 CC disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's  
 CC chorea, Down's syndrome, nerve deafness, Meniere's disease and also for  
 CC treating peripheral neuropathies, motorneuron disorders such as  
 CC amyotrophic lateral sclerosis (ALS), Bell's palsy and various conditions  
 CC involving spinal muscular atrophy and paralysis. The present sequence  
 CC is human presenilin PS1-FAD (familial Alzheimer's disease) mutant.  
 CC Note: This sequence is not shown in the specification but is derived from  
 CC human presenilin (PS1) protein [SEQ ID NO: 4] shown in page 60-61 of the  
 CC specification (AAE05466).

SQ Sequence 467 AA:

Query Match 56.7%; Score 265; DB 22; Length 467;  
 Best Local Similarity 99.6%; Pred. No. 5.3e-253;  
 Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MTELPAPLSTYFONAMSEDNHLSNTYVRSONDRERHNDRLSLGHPPLSLNGRPGNSR 60  
 |||||||  
 DB 1 MTELPAPLSTYFONAMSEDNHLSNTYVRSONDRERHNDRLSLGHPPLSLNGRPGNSR 60  
 QY 61 QVVEDEDEDELTLTKYGAHHYIMLFVPTLCMNVVVAITKVSFTTRKDGOLITPPE 120  
 |||||||  
 DB 61 QVVEDEDEDELTLTKYGAHHYIMLFVPTLCMNVVVAITKVSFTTRKDGOLITPPE 120  
 QY 121 DRETGORALHSILNAAIMISYIVVMTLLVLYKRCYKVHAWLIISSLLLEFFSFI 180  
 |||||||  
 DB 121 DRETGORALHSILNAAIMISYIVVMTLLVLYKRCYKVHAWLIISSLLLEFFSFI 180  
 QY 181 YLGEVFKTYNNAVDTYVALLIMNGVGMISIHMKGPLRLQOAVLIMISALMALVFIKY 240  
 |||||||  
 DB 181 YLGEVFKTYNNAVDTYVALLIMNGVGMISIHMKGPLRLQOAVLIMISALMALVFIKY 240  
 QY 241 LPBWTAMLILAVISYDLAVAVICPKGPLMLVETAOERNETLFPALITSTWMLVNMAE 300  
 |||||||  
 DB 241 LPBWTAMLILAVISYDLAVAVICPKGPLMLVETAOERNETLFPALITSTWMLVNMAE 300  
 QY 301 GDEPQORRVSKSKYNAESTERESODTYVAENDDGFSEMEQORSHLCPHRSTPESRA 360  
 |||||||  
 DB 301 GDEPQORRVSKSKYNAESTERESODTYVAENDDGFSEMEQORSHLCPHRSTPESRA 360  
 QY 361 VOELSSSILAGEDPERGVKLGDFIFYSVYVGKASATASGDWNTTACFAAILIGLCL 420  
 |||||||  
 DB 361 VOELSSSILAGEDPERGVKLGDFIFYSVYVGKASATASGDWNTTACFAAILIGLCL 420  
 QY 421 TLLLLAIFKKALPALPISTTGLVFFATDYLVOPMDLAFHOFYI 467  
 |||||||  
 DB 421 TLLLLAIFKKALPALPISTTGLVFFATDYLVOPMDLAFHOFYI 467

RESULT 48

AAM42375







QY 344 RDSHGGPHRSTPESRAAVQELSSSILAGEDPEERGVKLGDFIFYSVLVGKASATASGD 403  
 DB 342 RDSHGGPHRSTPESRAAVQELSSSILAGEDPEERGVKLGDFIFYSVLVGKASATASGD 401  
 QY 404 WNTTACFAVAILIGLCTLLLLAIFKKALPALPISITFGVLFYFTDYLVOPEMDQLAFH 463  
 DB 402 WNTTACFAVAILIGLCTLLLLAIFKKALPALPISITFGVLFYFTDYLVOPEMDQLAFH 461  
 QY 464 OFYI 467  
 DB 462 OFYI 465  
 RESULT 50  
 AAM12376  
 ID AAM12376 standard; protein: 463 AA.  
 AC AAM12376;  
 DT 17-JUN-1997 (first entry)  
 DE Human S182 gene product.  
 KW S182 gene; familial Alzheimer's disease; diagnosis;  
 KW transgenic animal.  
 OS Homo sapiens.  
 PN W09703999-A1.  
 PD 06-FEB-1997.  
 PE 26-JUN-1996; 96WO-US11065.  
 PR 02-AUG-1995; 95US-0001800.  
 PR 18-JUL-1995; 95US-0001500.  
 PA (UNSW) UNIV SOUTH FLORIDA.  
 PI Goale AM, Hardy JA;  
 DR WPI: 1997-132571/12.  
 DR N-PSDB: AAT63207.  
 PT New mutants of the S182 gene associated with familial Alzheimer's  
 PT disease - and related protein and transgenic animals, useful as  
 PT models for screening and assessing potential drugs  
 PS Disclosure; Fig 1A-D; 26pp; English.  
 CC A polypeptide (AAM12376) is the product of an S182 gene cDNA clone  
 CC (AAT63207) isolated from a human brain library. Several mutations in  
 CC the S182 gene have been found in families with members affected by  
 CC early onset Alzheimer's disease (AD): in 2 families Met to Val at  
 CC position 135; in 3 families Met to Val at 142; in 1 family Pro to  
 CC Ser at 263; in 4 families Glu to Ala at 276; and in 1 family Glu to  
 CC Gly at 280. Detection of the mutations is used to diagnose AD, or  
 CC a predisposition to it. Transgenic animals can be produced that  
 CC are useful as models for screening and assessing potential drugs.  
 SQ Sequence 463 AA:  
 Query Match 50.5%; Score 236; DB 18; Length 463;  
 Best Local Similarity 99.5%; Pred. No. 2.4e-224;  
 Matches 436; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 30 NDNRERQNDNRSLGHPPLSGNRPQNSROVDEDEDELTLKYGAKHVMLEFVPV 89  
 DB 26 NDNRERQNDNRSLGHPPLSGNRPQNSROVDEDEDELTLKYGAKHVMLEFVPV 85  
 QY 90 TLCAVVVAVTIKSVSFYTRKDGQLIYPTFEDTETVGORALHSILNAAIMISIVVMTIL 149  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 86 TLCAVVVAVTIKSVSFYTRKDGQLIYPTFEDTETVGORALHSILNAAIMISIVVMTIL 145  
 QY 150 LVVLYKYRCYKVIHAWLLISSLILFFPSFTYLGVEFTYNNVAVDYTVALLIHWLVYG 209  
 DB 146 LVVLYKYRCYKVIHAWLLISSLILFFPSFTYLGVEFTYNNVAVDYTVALLIHWLVYG 205  
 QY 210 MISIMKGPPLRQOAYLIMISALMALVFTKYLPEWTAMLLAVISVYDLVAVLCRKPRLR 269  
 DB 206 MISIMKGPPLRQOAYLIMISALMALVFTKYLPEWTAMLLAVISVYDLVAVLCRKPRLR 265  
 QY 270 MLVETAQRNETTLFPALIVSSSTMVWLVMNAEGDPEAQRVSKNSKYNAESTERESQDTVA 329  
 DB 266 MLVETAQRNETTLFPALIVSSSTMVWLVMNAEGDPEAQRVSKNSKYNAESTERESQDTVA 325  
 QY 330 EMDDGGSFEWEAQRDSSHGPHRSTPESRAAVQELSSSILAGEDPEERGVKLGDFIFY 389  
 DB 326 EMDDGGSFEWEAQRDSSHGPHRSTPESRAAVQELSSSILAGEDPEERGVKLGDFIFY 385  
 QY 390 SVLVGKASATASGDMNTTIACFVAILIGLCTLLLLAIFKKALPALPISITFGVLFYFAT 449  
 DB 386 SVLVGKASATASGDMNTTIACFVAILIGLCTLLLLAIFKKALPALPISITFGVLFYFAT 445  
 QY 450 DYLVOPEMDQLAFHOFYI 467  
 DB 446 DYLVOPEMDQLAFHOFYI 463

Search completed: February 6, 2003, 14:18:59  
 Job time : 86.0394 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: February 6, 2003, 14:17:07 : Search time 51.3506 Seconds  
(without alignments)  
1873.862 Million cell updates/sec

Title: US-09-689-159a-2

Sequence: 1 MFELPAPLSTYFNAQMSDN.....ATDYLVQPFMDLAFHQFYI 467

## Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

## Database :

SPREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_podent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriapi:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Query	Length	DB	ID	Description
1	204	43.7	289	4	Q9UIF0	Q9UIF0 homo sapien
2	161	34.5	184	4	Q95465	Q95465 homo sapien
3	127	27.2	378	4	Q96P33	Q96P33 homo sapien
4	87	18.6	261	1	Q91WK6	Q91WK6 mus musculu
5	63	13.5	468	13	Q90X08	Q90X08 gallus gall
6	59	12.6	384	13	Q73869	Q73869 cyprinus ca
7	42	9.0	456	13	Q9W6T7	Q9W6T7 brachydanto
8	39	8.4	441	13	Q91991	Q91991 brachydanto
9	39	8.4	441	13	Q902E4	Q902E4 brachydanto
10	35	7.5	504	5	Q8MS58	Q8MS58 branchiosto
11	35	7.5	525	5	Q8MS59	Q8MS59 branchiosto
12	28	6.0	332	11	Q9D616	Q9D616 mus musculu
13	28	6.0	390	4	Q96P32	Q96P32 homo sapien
14	28	6.0	448	11	Q91VS3	Q91VS3 mus musculu
15	23	4.9	451	13	Q90X07	Q90X07 gallus gall
16	20	4.3	582	5	Q9GJ38	Q9GJ38 helix lucor

17	3.2	397	10	Q9S1K7	Q9S1K7 arabidopsis
18	3.2	478	10	Q9SDG2	Q9SDG2 oryza sativ
19	1.9	334	17	Q97UC3	Q97UC3 sulfolobus
20	1.7	64	13	Q9PV06	Q9PV06 pleocoglossu
21	1.7	67	13	Q9PV05	Q9PV05 harpador sp
22	1.7	83	2	Q93EP9	Q93EP9 helicobacte
23	1.7	104	12	Q8ORD1	Q8ORD1 white spot
24	1.7	105	12	Q8VAV0	Q8VAV0 white spot
25	1.7	139	11	Q63071	Q63071 ratius norv
26	1.7	288	2	Q8VQL7	Q8VQL7 bruceia ab
27	1.7	314	10	Q9SLP4	Q9SLP4 tulipa bake
28	1.7	317	16	Q8ZFL5	Q8ZFL5 yerania pe
29	1.7	347	16	Q9RR94	Q9RR94 deinococcus
30	1.7	354	3	Q60121	Q60121 schizosacch
31	1.7	369	6	Q7794	Q7794 canis famli
32	1.7	379	5	Q966H2	Q966H2 caenorhabdi
33	1.7	392	10	Q92PX3	Q92PX3 arabidopsis
34	1.7	519	5	Q9V4U2	Q9V4U2 drosophila
35	1.7	522	2	Q52557	Q52557 amycolators
36	1.7	538	11	Q03982	Q03982 cricetus
37	1.7	563	11	Q8R569	Q8R569 mus musculu
38	1.7	581	16	Q97178	Q97178 clostridium
39	1.7	628	4	Q95052	Q95052 homo sapien
40	1.7	668	6	Q8WNU6	Q8WNU6 canis famli
41	1.7	749	11	Q8RUK1	Q8RUK1 mus musculu
42	1.7	1022	11	Q63928	Q63928 mus sp. big
43	1.7	1163	6	Q9TSU2	Q9TSU2 felis silve
44	1.7	1169	11	Q60502	Q60502 cricetus
45	1.7	1231	5	Q97140	Q97140 dictyosteli
46	1.7	1272	11	Q9UK64	Q9UK64 ratius norv
47	1.7	1272	11	Q924L6	Q924L6 ratius norv
48	1.7	1280	6	Q46605	Q46605 canis famli
49	1.7	1285	6	Q02793	Q02793 ovis aries
50	1.7	1726	5	Q02569	Q02569 plasmodium
51	1.7	2958	5	Q9VDY5	Q9VDY5 drosophila
52	1.5	29	15	Q905H6	Q905H6 human immun
53	1.5	19	8	Q9G355	Q9G355 trapezole sa
54	1.5	45	16	Q8U5R7	Q8U5R7 agrobacteri
55	1.5	62	2	Q50445	Q50445 mycobacteri
56	1.5	62	13	Q90WP7	Q90WP7 rana pipien
57	1.5	70	13	Q8QFQ3	Q8QFQ3 rana pipien
58	1.5	71	13	Q13103	Q13103 cyprinus ca
59	1.5	78	16	Q9ZNM05	Q9ZNM05 helicobacte
60	1.5	86	13	Q91RX7	Q91RX7 petromyzon
61	1.5	86	16	Q84166	Q84166 chlamydia t
62	1.5	87	13	Q918X9	Q918X9 petromyzon
63	1.5	90	17	Q8U0W3	Q8U0W3 pyrococcus
64	1.5	91	10	Q9AS79	Q9AS79 oryza sativ
65	1.5	94	2	Q8VLS5	Q8VLS5 enterococu
66	1.5	95	4	Q8WYR9	Q8WYR9 homo sapien
67	1.5	95	13	Q91BX8	Q91BX8 petromyzon
68	1.5	97	11	Q9ERU0	Q9ERU0 mus musculu
69	1.5	99	11	Q9ERU1	Q9ERU1 ratius norv
70	1.5	103	16	Q8XEP2	Q8XEP2 salmonella
71	1.5	104	16	Q9HYV9	Q9HYV9 pseudomonas
72	1.5	105	2	Q9EYX3	Q9EYX3 klebsiella
73	1.5	107	17	Q9YAQ5	Q9YAQ5 aeropyrum p
74	1.5	108	17	Q9YC54	Q9YC54 aeropyrum p
75	1.5	109	17	Q9YG06	Q9YG06 aeropyrum p
76	1.5	115	12	Q65536	Q65536 baboon herp
77	1.5	117	3	Q96U14	Q96U14 neurospora
78	1.5	125	11	Q9D9E2	Q9D9E2 mus musculu
79	1.5	125	16	Q8U0J9	Q8U0J9 agrobacteri
80	1.5	131	16	Q9A6X3	Q9A6X3 caulobacter
81	1.5	132	16	Q8XHT7	Q8XHT7 clostridium
82	1.5	139	2	Q9X541	Q9X541 corynebacte
83	1.5	139	5	Q9V556	Q9V556 drosophila
84	1.5	149	11	Q35327	Q35327 mus musculu
85	1.5	149	16	Q8R7B1	Q8R7B1 thermotanaer
86	1.5	155	8	Q9MN89	Q9MN89 galacthealin
87	1.5	155	11	Q9DB92	Q9DB92 mus musculu
88	1.5	157	10	Q8RV90	Q8RV90 oryza sativ
89	1.5	157	12	Q8VAG6	Q8VAG6 white spot

90	7	1.5	168	12	036423	036423 alcelaphine	163	7	1.5	299	10	092VK1	092VK1 arabisdopsis
91	7	1.5	169	17	09YC98	09YC98 aeropyrum p	164	7	1.5	310	10	09LRX5	09LRX5 arabisdopsis
92	7	1.5	175	12	08Q587	08Q587 chimpanzee	165	7	1.5	311	12	08V719	08V719 simian herp
93	7	1.5	176	16	09KJC4	09KJC4 bacillus ha	166	7	1.5	312	8	09MFM8	09MFM8 cochlomyia
94	7	1.5	178	5	09KZJ1	09KZJ1 drosophila	167	7	1.5	314	8	09MM08	09MM08 anthoceros
95	7	1.5	178	17	097C37	097C37 thermoplasma	168	7	1.5	315	17	08TLX3	08TLX3 methanosarc
96	7	1.5	179	5	09NAC5	09NAC5 caenorhabd	169	7	1.5	316	11	08R2A2	08R2A2 mus musculu
97	7	1.5	180	8	08SI08	08SI08 dekenia im	170	7	1.5	317	16	08U848	08U848 agrobacteri
98	7	1.5	183	16	08XTM6	08XTM6 raistonia s	171	7	1.5	318	8	099805	099805 ixodes hexa
99	7	1.5	184	10	022258	022258 arabisdopsis	172	7	1.5	319	8	09GPF5	09GPF5 dirosyrea wint
100	7	1.5	192	5	09N7N6	09N7N6 leishmania	173	7	1.5	321	8	09GPF8	09GPF8 dioscorea b
101	7	1.5	194	11	09IXB5	09IXB5 mus musculu	174	7	1.5	323	8	09B635	09B635 anguilla ja
102	7	1.5	194	17	09YFK1	09YFK1 aeropyrum p	175	7	1.5	323	8	09GFG8	09GFG8 asarum cana
103	7	1.5	195	2	09FOP4	09FOP4 listeria mo	176	7	1.5	324	16	034616	034616 bacillus su
104	7	1.5	201	4	09UM05	09UM05 homo sapien	177	7	1.5	324	8	09GPF0	09GPF0 illicium pa
105	7	1.5	210	13	091596	091596 xenopus lae	178	7	1.5	324	8	09GPF6	09GPF6 litiodendro
106	7	1.5	210	16	09KAP8	09KAP8 bacillus ha	179	7	1.5	324	8	09GPF2	09GPF2 trochodendr
107	7	1.5	211	16	099R14	099R14 staphylococ	180	7	1.5	325	6	09BR43	09BR43 felis silve
108	7	1.5	212	11	09CRJ0	09CRJ0 mus musculu	181	7	1.5	325	6	09BR42	09BR42 leopardus p
109	7	1.5	213	16	08XN92	08XN92 clostridium	182	7	1.5	325	6	09BR41	09BR41 panthera on
110	7	1.5	216	8	09TAG4	09TAG4 uma exsul.	183	7	1.5	325	8	09GFM0	09GFM0 acorus cala
111	7	1.5	216	9	08W628	08W628 bacterioph	184	7	1.5	325	8	09GFG2	09GFG2 ceratophyll
112	7	1.5	216	13	090228	090228 ambystoma m	185	7	1.5	325	8	09GFG0	09GFG0 cercidiphyll
113	7	1.5	217	8	09TAG5	09TAG5 uma inornat	186	7	1.5	325	8	09GFE8	09GFE8 lactoris fe
114	7	1.5	217	8	09T3Y7	09T3Y7 uma notata.	187	7	1.5	325	8	09GFE4	09GFE4 saurus ce
115	7	1.5	217	13	090WM4	090WM4 xenopus lae	188	7	1.5	325	13	09PSX3	09PSX3 xenopus lae
116	7	1.5	218	16	092BS8	092BS8 listeria in	189	7	1.5	326	6	09BRF0	09BRF0 lemur catra
117	7	1.5	218	10	09KMT8	09KMT8 listeria mo	190	7	1.5	326	6	09BRF4	09BRF4 sus scrofa
118	7	1.5	219	10	09FOZ8	09FOZ8 nicotiana t	191	7	1.5	326	8	09GEG6	09GEG6 cabomba car
119	7	1.5	221	10	08S032	08S032 oryza sativ	192	7	1.5	326	8	09GEG4	09GEG4 calycanthus
120	7	1.5	226	12	072839	072839 porcine epi	193	7	1.5	326	8	09GFF4	09GFF4 ginkgo billo
121	7	1.5	226	12	091AU9	091AU9 porcine epi	194	7	1.5	326	8	09MSQ7	09MSQ7 zania furfu
122	7	1.5	226	12	037351	037351 porcine epi	195	7	1.5	326	8	09MSP5	09MSP5 nympheaa od
123	7	1.5	226	12	007504	007504 porcine epi	196	7	1.5	326	8	09GCE32	09GCE32 amborella t
124	7	1.5	228	11	064057	064057 retus sp.	197	7	1.5	326	11	099NR4	099NR4 hystrix dra
125	7	1.5	232	5	095RA2	095RA2 drosophila	198	7	1.5	326	11	099NR3	099NR3 erehthizon d
126	7	1.5	235	11	08VH42	08VH42 sigmodon hi	199	7	1.5	326	11	099NR1	099NR1 heteroceph
127	7	1.5	236	5	09GCV9	09GCV9 caenorhabd	200	7	1.5	327	16	09A5T8	09A5T8 caluobacter
128	7	1.5	237	8	08WED7	08WED7 pameleaesci	201	7	1.5	330	10	09LMJ9	09LMJ9 arabisdopsis
129	7	1.5	238	1	09VPK4	09VPK4 drosophila	202	7	1.5	333	8	09BR08	09BR08 gullardia
130	7	1.5	241	2	09JN88	09JN88 streptomyc	203	7	1.5	337	13	042344	042344 cynops pyr
131	7	1.5	241	11	062076	062076 mus musculu	204	7	1.5	338	16	08Y0B2	08Y0B2 raiistonia s
132	7	1.5	243	4	09MTB3	09MTB3 homo sapien	205	7	1.5	340	5	09G7C9	09G7C9 dictyostell
133	7	1.5	252	6	09BRF40	09BRF40 canis famli	206	7	1.5	341	5	09GSO1	09GSO1 caenorhabd
134	7	1.5	252	10	09SCH24	09SCH24 arabisdopsis	207	7	1.5	342	10	P93153	P93153 gossypium h
135	7	1.5	252	10	09C811	09C811 arabisdopsis	208	7	1.5	343	8	09B1J5	09B1J5 uma inornat
136	7	1.5	257	5	08SMV1	08SMV1 drosophila	209	7	1.5	343	8	09B1J4	09B1J4 uma notata.
137	7	1.5	257	11	09D1H9	09D1H9 mus musculu	210	7	1.5	343	8	09B1J3	09B1J3 uma notata
138	7	1.5	257	17	09HJTO	09HJTO thermoplas	211	7	1.5	343	8	09B0S7	09B0S7 uma notata
139	7	1.5	259	16	08RCM5	08RCM5 thermoaear	212	7	1.5	343	8	09B4I1	09B4I1 uma inornat
140	7	1.5	259	17	09V206	09V206 pyrococcus	213	7	1.5	343	8	09B4I0	09B4I0 uma notata
141	7	1.5	262	16	09KDO1	09KDO1 bacillus ha	214	7	1.5	343	8	09B4H9	09B4H9 uma notata.
142	7	1.5	264	13	09YGE1	09YGE1 oncorhynch	215	7	1.5	343	8	09B4H7	09B4H7 uma notata.
143	7	1.5	265	5	017895	017895 caenorhabd	216	7	1.5	343	8	09B4H5	09B4H5 uma scopari
144	7	1.5	267	13	09YGE2	09YGE2 oncorhynch	217	7	1.5	343	16	09BAV4	09BAV4 rhizobium m
145	7	1.5	269	2	09FAH1	09FAH1 pseudomonas	218	7	1.5	343	16	092VZ0	092VZ0 rhizobium l
146	7	1.5	269	16	09ACS2	09ACS2 streptomyc	219	7	1.5	344	4	013079	013079 homo sapien
147	7	1.5	274	5	022843	022843 caenorhabd	220	7	1.5	346	5	09NLA3	09NLA3 asterina pe
148	7	1.5	275	5	09BPN1	09BPN1 caenorhabd	221	7	1.5	346	5	08WSV9	08WSV9 asterina pe
149	7	1.5	276	4	09H913	09H913 homo sapien	222	7	1.5	346	8	0954E1	0954E1 pseudochelr
150	7	1.5	278	8	09TNJ8	09TNJ8 sphagnum fa	223	7	1.5	347	8	021636	021636 darevskia c
151	7	1.5	280	8	09TNJ3	09TNJ3 haplomitriu	224	7	1.5	347	8	021647	021647 darevskia c
152	7	1.5	283	4	09BOJ0	09BOJ0 homo sapien	225	7	1.5	347	8	09MNY7	09MNY7 darevskia m
153	7	1.5	285	5	09U2H9	09U2H9 caenorhabd	226	7	1.5	347	8	09MNS3	09MNS3 lacerta nai
154	7	1.5	286	17	08RT83	08RT83 methanosarc	227	7	1.5	347	8	0954E0	0954E0 trichosurus
155	7	1.5	290	8	09TNL2	09TNL2 physcomitre	228	7	1.5	347	8	0952S5	0952S5 trichosurus
156	7	1.5	291	2	093TU0	093TU0 acidithioba	229	7	1.5	351	5	08T3I8	08T3I8 drosophila
157	7	1.5	294	10	094AK1	094AK1 arabisdopsis	230	7	1.5	351	16	08XPA5	08XPA5 anabaena sp
158	7	1.5	295	10	09LSZ8	09LSZ8 arabisdopsis	231	7	1.5	353	8	09BBT0	09BBT0 lotus japon
159	7	1.5	296	16	09LA44	09LA44 shigella dy	232	7	1.5	353	8	08W119	08W119 psilocium nu
160	7	1.5	296	16	08X7F2	08X7F2 escherichia	233	7	1.5	353	10	08S8X8	08S8X8 atropa bell
161	7	1.5	298	2	09REN4	09REN4 zymomonas m	234	7	1.5	355	17	097WJ7	097WJ7 sulfolobus
162	7	1.5	299	8	09TN18	09TN18 adiantum ca	235	7	1.5	356	13	09DE33	09DE33 xenopus lae

236	7	1.5	357	16	Q8YU25	Q8YU25 anabaena sp	309	7	1.5	479	17	Q27332	Q27332 methanobact
237	7	1.5	358	16	Q8YIP7	Q8YIP7 ralstonia s	310	7	1.5	482	10	Q9FPT4	Q9FPT4 arabidopsis
238	7	1.5	360	16	Q8U6M3	Q8U6M3 agrobacteri	311	7	1.5	482	10	Q949Y0	Q949Y0 arabidopsis
239	7	1.5	361	4	Q96C80	Q96C80 homo sapien	312	7	1.5	480	10	Q9C8H9	Q9C8H9 arabidopsis
240	7	1.5	361	11	Q91Y81	Q91Y81 rattus norv	313	7	1.5	483	5	Q17554	Q17554 caenorhabdi
241	7	1.5	364	3	Q9USL0	Q9USL0 schizosacch	314	7	1.5	484	10	Q9LHF1	Q9LHF1 arabidopsis
242	7	1.5	364	5	Q9VK91	Q9VK91 drosophila	315	7	1.5	494	12	Q90JL1	Q90JL1 cercopithec
243	7	1.5	365	16	Q8RBB7	Q8RBB7 thermomanae	316	7	1.5	494	17	Q8U246	Q8U246 pyrococcus
244	7	1.5	366	2	Q8VO93	Q8VO93 uncultured	317	7	1.5	497	16	Q92X23	Q92X23 rhizobium m
245	7	1.5	367	16	Q8ZDY6	Q8ZDY6 yersinia pe	318	7	1.5	498	12	Q91ML7	Q91ML7 lumpy skin
246	7	1.5	368	10	Q9SMR6	Q9SMR6 arabidopsis	319	7	1.5	498	16	Q92U87	Q92U87 rhizobium m
247	7	1.5	371	3	Q9HFR1	Q9HFR1 cryptococcu	320	7	1.5	499	12	Q913H9	Q913H9 pseudorabie
248	7	1.5	373	12	Q8U2J8	Q8U2J8 cercopithec	321	7	1.5	500	10	Q940M4	Q940M4 arabidopsis
249	7	1.5	375	17	Q57876	Q57876 pyrococcus	322	7	1.5	502	8	Q9MGK6	Q9MGK6 chrysodidym
250	7	1.5	376	16	Q926X9	Q926X9 listeria in	323	7	1.5	508	4	Q15208	Q15208 homo sapien
251	7	1.5	376	16	Q8Y3R5	Q8Y3R5 listeria in	324	7	1.5	514	16	Q930U7	Q930U7 rhizobium m
252	7	1.5	377	5	Q17446	Q17446 caenorhabdi	325	7	1.5	515	2	P96182	P96182 wolfinella s
253	7	1.5	378	4	Q9BWI3	Q9BWI3 homo sapien	326	7	1.5	517	5	Q9U266	Q9U266 caenorhabdi
254	7	1.5	379	8	Q9B6F0	Q9B6F0 thryonomy	327	7	1.5	520	5	Q9GNK5	Q9GNK5 drosophila
255	7	1.5	379	16	Q983P6	Q983P6 rhizobium l	328	7	1.5	522	16	Q9RUT4	Q9RUT4 delnoccocus
256	7	1.5	379	17	Q9YFW4	Q9YFW4 aetopyrum p	329	7	1.5	527	4	Q14706	Q14706 homo sapien
257	7	1.5	381	8	Q63532	Q63532 antechinus	330	7	1.5	531	10	Q9FEP8	Q9FEP8 arabidopsis
258	7	1.5	381	8	Q9AM96	Q9AM96 Isodon mac	331	7	1.5	534	13	Q93296	Q93296 gallus gall
259	7	1.5	381	8	Q35545	Q35545 planigale s	332	7	1.5	535	16	Q92L32	Q92L32 rhizobium m
260	7	1.5	385	3	Q9P981	Q9P981 robillarda	333	7	1.5	537	10	Q9S638	Q9S638 arabidopsis
261	7	1.5	385	16	Q914M7	Q914M7 pseudomonas	334	7	1.5	545	5	Q9VMT4	Q9VMT4 drosophila
262	7	1.5	385	16	Q98F78	Q98F78 rhizobium l	335	7	1.5	555	16	Q99PTD2	Q99PTD2 staphylococ
263	7	1.5	387	10	Q942S1	Q942S1 oryza sativ	336	7	1.5	561	5	Q960U5	Q960U5 drosophila
264	7	1.5	388	12	Q65360	Q65360 oryza pseu	337	7	1.5	566	4	Q9UP68	Q9UP68 homo sapien
265	7	1.5	391	5	Q9N6C3	Q9N6C3 caenorhabdi	338	7	1.5	566	11	Q9ERQ3	Q9ERQ3 mus musculu
266	7	1.5	392	16	Q67276	Q67276 aquifex aeo	339	7	1.5	569	13	Q9PYL1	Q9PYL1 gallus gall
267	7	1.5	393	5	Q967G9	Q967G9 chironomus	340	7	1.5	579	4	Q43386	Q43386 homo sapien
268	7	1.5	395	5	Q95SS2	Q95SS2 drosophila	341	7	1.5	579	4	Q96J15	Q96J15 homo sapien
269	7	1.5	396	13	Q73669	Q73669 gallus gall	342	7	1.5	580	5	Q9VDV5	Q9VDV5 drosophila
270	7	1.5	397	10	Q9FHB1	Q9FHB1 arabidopsis	343	7	1.5	581	5	Q22620	Q22620 caenorhabdi
271	7	1.5	397	16	Q54177	Q54177 streptomyc	344	7	1.5	585	5	Q8SX51	Q8SX51 drosophila
272	7	1.5	398	4	Q9H3Q3	Q9H3Q3 homo sapien	345	7	1.5	585	12	Q89244	Q89244 woodchuck h
273	7	1.5	401	10	Q9LXU7	Q9LXU7 arabidopsis	346	7	1.5	596	10	Q9LSG6	Q9LSG6 arabidopsis
274	7	1.5	402	4	Q8WVX6	Q8WVX6 homo sapien	347	7	1.5	605	13	Q9M6Q9	Q9M6Q9 fugu rubrip
275	7	1.5	405	11	Q91W19	Q91W19 mus musculu	348	7	1.5	610	16	Q9K6U0	Q9K6U0 bacillus ha
276	7	1.5	406	3	Q12161	Q12161 saccharomyc	349	7	1.5	611	5	Q9VSV9	Q9VSV9 drosophila
277	7	1.5	416	12	Q9YPA9	Q9YPA9 kaposi's sa	350	7	1.5	616	4	Q96PK3	Q96PK3 homo sapien
278	7	1.5	417	4	Q8WXX6	Q8WXX6 homo sapien	351	7	1.5	622	11	Q923C5	Q923C5 mus musculu
279	7	1.5	417	5	Q16377	Q16377 caenorhabdi	352	7	1.5	622	16	Q92BP2	Q92BP2 rhizobium m
280	7	1.5	425	3	Q12700	Q12700 debaromyce	353	7	1.5	624	5	Q9WZT0	Q9WZT0 drosophila
281	7	1.5	425	5	Q9VUT8	Q9VUT8 drosophila	354	7	1.5	641	5	Q8T5O6	Q8T5O6 heterodera
282	7	1.5	427	2	Q9ZSD5	Q9ZSD5 rhodobacter	355	7	1.5	643	4	Q96E81	Q96E81 homo sapien
283	7	1.5	428	10	Q8S3Q9	Q8S3Q9 oryza sativ	356	7	1.5	652	17	Q976X0	Q976X0 sulfolobus
284	7	1.5	428	8	Q950U6	Q950U6 hyaloraphid	357	7	1.5	657	4	Q8TAM1	Q8TAM1 homo sapien
285	7	1.5	429	2	Q8RNN2	Q8RNN2 legionella	358	7	1.5	671	10	Q9FVB9	Q9FVB9 arabidopsis
286	7	1.5	429	17	Q8RPG9	Q8RPG9 methanosarc	359	7	1.5	682	5	Q9GTV3	Q9GTV3 plasmodium
287	7	1.5	432	5	Q9TXH6	Q9TXH6 caenorhabdi	360	7	1.5	684	10	Q91LX1	Q91LX1 arabidopsis
288	7	1.5	434	4	Q9BS03	Q9BS03 homo sapien	361	7	1.5	685	5	Q17378	Q17378 caenorhabdi
289	7	1.5	438	16	Q8XEV7	Q8XEV7 salmonella	362	7	1.5	685	5	Q20423	Q20423 caenorhabdi
290	7	1.5	440	16	Q9KLL9	Q9KLL9 vibrio chol	363	7	1.5	687	5	Q94960	Q94960 drosophila
291	7	1.5	444	5	Q9XVA2	Q9XVA2 caenorhabdi	364	7	1.5	687	16	Q07753	Q07753 mycobacteri
292	7	1.5	444	11	Q9D5R4	Q9D5R4 mus musculu	365	7	1.5	689	10	Q65312	Q65312 arabidopsis
293	7	1.5	446	10	Q94C58	Q94C58 arabidopsis	366	7	1.5	692	2	Q33448	Q33448 proteus mir
294	7	1.5	448	5	Q9NFE7	Q9NFE7 echinococcu	367	7	1.5	695	13	Q9D6J8	Q9D6J8 gallus gall
295	7	1.5	449	4	Q9H6S4	Q9H6S4 homo sapien	368	7	1.5	702	4	Q9HOK9	Q9HOK9 homo sapien
296	7	1.5	449	16	Q9ABE3	Q9ABE3 caulobacter	369	7	1.5	712	11	Q54729	Q54729 rattus norv
297	7	1.5	452	10	Q9ZQO0	Q9ZQO0 arabidopsis	370	7	1.5	714	2	Q8RPL9	Q8RPL9 ehrlchia c
298	7	1.5	453	4	Q9NW27	Q9NW27 homo sapien	371	7	1.5	715	11	Q9Q2X1	Q9Q2X1 rattus norv
299	7	1.5	460	10	Q23014	Q23014 arabidopsis	372	7	1.5	722	5	Q9UAB8	Q9UAB8 leishmania
300	7	1.5	460	10	Q9SH79	Q9SH79 arabidopsis	373	7	1.5	731	10	Q9EVB8	Q9EVB8 ureaplasma
301	7	1.5	461	2	Q51524	Q51524 pseudomonas	374	7	1.5	731	16	Q9POU4	Q9POU4 oryza sativ
302	7	1.5	462	10	Q38698	Q38698 asarum euro	375	7	1.5	738	10	Q8S6U7	Q8S6U7 oryza sativ
303	7	1.5	462	11	Q923F4	Q923F4 mus musculu	376	7	1.5	738	11	Q9OXO2	Q9OXO2 rattus norv
304	7	1.5	464	3	Q8WZM0	Q8WZM0 yarrowia li	377	7	1.5	740	4	Q14747	Q14747 homo sapien
305	7	1.5	465	3	Q06525	Q06525 saccharomyc	378	7	1.5	740	4	Q9UER7	Q9UER7 homo sapien
306	7	1.5	470	10	Q9LSH7	Q9LSH7 arabidopsis	379	7	1.5	740	4	Q15141	Q15141 homo sapien
307	7	1.5	474	10	Q8RY61	Q8RY61 arabidopsis	380	7	1.5	743	5	Q8SX92	Q8SX92 drosophila
308	7	1.5	478	5	Q43955	Q43955 leishmania	381	7	1.5	746	11	Q63553	Q63553 rattus norv

382	7	1.5	748	5	Q9W091	Q9W091 drosophila	455	7	1.5	1237	16	Q8YH6	Q8YH6 anabaena sp
383	7	1.5	748	11	Q91WX6	Q91WX6 mus musculus	456	7	1.5	1245	5	Q96195	Q96195 plasmodium
384	7	1.5	748	11	Q8VDU5	Q8VDU5 mus musculus	457	7	1.5	1254	10	Q9SR21	Q9SR21 arabidopsis
385	7	1.5	748	16	Q8Z411	Q8Z411 salmonella	458	7	1.5	1262	5	Q61771	Q61771 caenorhabd
386	7	1.5	748	16	Q8X6K3	Q8X6K3 escherichia	459	7	1.5	1287	4	Q92549	Q92549 homo sapien
387	7	1.5	751	3	Q9P749	Q9P749 neotropora	460	7	1.5	1288	13	Q93437	Q93437 gallus gall
388	7	1.5	751	13	Q9DGJ7	Q9DGJ7 gallus gall	461	7	1.5	1291	11	Q61812	Q61812 mus musculus
389	7	1.5	765	13	Q9NRH2	Q9NRH2 homo sapien	462	7	1.5	1317	16	Q91181	Q91181 pseudomonas
390	7	1.5	775	10	Q8RWL9	Q8RWL9 fundulus he	463	7	1.5	1342	10	Q9FKN5	Q9FKN5 arabidopsis
391	7	1.5	775	10	Q8RWL9	Q8RWL9 arabidopsis	464	7	1.5	1343	11	Q64730	Q64730 mus musculus
392	7	1.5	778	16	Q8UCO3	Q8UCO3 agrobacteri	465	7	1.5	1391	4	Q8TEU6	Q8TEU6 homo sapien
393	7	1.5	782	5	Q9GV92	Q9GV92 caenorhabd	466	7	1.5	1407	10	Q80635	Q80635 arabidopsis
394	7	1.5	789	2	Q93M73	Q93M73 caenorhabd	467	7	1.5	1432	16	Q935P0	Q935P0 salmonella
395	7	1.5	790	5	Q01264	Q01264 xanthomonas	468	7	1.5	1457	4	Q96PS7	Q96PS7 homo sapien
396	7	1.5	796	13	Q8UW36	Q8UW36 fluu rubrip	469	7	1.5	1457	11	Q920J0	Q920J0 mus musculus
397	7	1.5	799	4	Q94874	Q94874 homo sapien	470	7	1.5	1457	11	Q920I8	Q920I8 rattus norv
398	7	1.5	808	16	Q92SS9	Q92SS9 rhizobium m	471	7	1.5	1469	10	Q9LKR1	Q9LKR1 pisum sativ
399	7	1.5	809	16	Q8ZCB5	Q8ZCB5 yersinia pe	472	7	1.5	1480	4	Q96G04	Q96G04 homo sapien
400	7	1.5	815	5	Q44385	Q44385 caenorhabd	473	7	1.5	1484	9	Q96G00	Q96G00 roseophag
401	7	1.5	817	4	Q9BRK2	Q9BRK2 homo sapien	474	7	1.5	1488	11	Q9ERH3	Q9ERH3 rattus norv
402	7	1.5	820	11	Q922R5	Q922R5 mus musculu	475	7	1.5	1489	11	Q920I9	Q920I9 mus musculu
403	7	1.5	820	16	Q92TA0	Q92TA0 rhizobium m	476	7	1.5	1509	4	Q96PC1	Q96PC1 homo sapien
404	7	1.5	822	10	Q41370	Q41370 spinacia ol	477	7	1.5	1508	3	Q74415	Q74415 schizosacch
405	7	1.5	851	5	Q46093	Q46093 drosophila	478	7	1.5	1596	5	Q23178	Q23178 caenorhabd
406	7	1.5	879	10	Q41010	Q41010 pisum sativ	479	7	1.5	1601	4	Q8TEU7	Q8TEU7 homo sapien
407	7	1.5	882	12	Q918N6	Q918N6 woodchuck h	480	7	1.5	1621	5	Q8TEJ5	Q8TEJ5 dictyostell
408	7	1.5	884	12	Q918N8	Q918N8 woodchuck h	481	7	1.5	1644	10	Q9ZOK0	Q9ZOK0 arabidopsis
409	7	1.5	884	12	Q918N8	Q918N8 woodchuck h	482	7	1.5	1687	5	Q18259	Q18259 caenorhabd
410	7	1.5	884	12	Q918N7	Q918N7 woodchuck h	483	7	1.5	1761	10	Q48867	Q48867 daucus caro
411	7	1.5	884	12	Q918N7	Q918N7 woodchuck h	484	7	1.5	1905	5	Q9Y4V8	Q9Y4V8 drosophila
412	7	1.5	884	12	Q918N5	Q918N5 woodchuck h	485	7	1.5	1972	4	Q9ULF8	Q9ULF8 homo sapien
413	7	1.5	884	12	Q918N4	Q918N4 woodchuck h	486	7	1.5	2047	4	Q15019	Q15019 homo sapien
414	7	1.5	884	12	Q918N3	Q918N3 woodchuck h	487	7	1.5	2078	5	Q8BPN1	Q8BPN1 oikopleura
415	7	1.5	904	16	Q9HUM7	Q9HUM7 pseudomonas	488	7	1.5	2102	5	Q8RTP7	Q8RTP7 dictyostell
416	7	1.5	910	12	Q9G9T5	Q9G9T5 avian adeno	489	7	1.5	2157	3	Q00333	Q00333 coprinus cl
417	7	1.5	914	5	Q16530	Q16530 caenorhabd	490	7	1.5	2276	4	Q75050	Q75050 homo sapien
418	7	1.5	920	5	Q9N306	Q9N306 caenorhabd	491	7	1.5	2380	5	Q96266	Q96266 plasmodium
419	7	1.5	938	5	Q9V3X5	Q9V3X5 drosophila	492	7	1.5	2410	5	Q8SR14	Q8SR14 encephalit
420	7	1.5	939	16	Q9PLB7	Q9PLB7 chlamydia m	493	7	1.5	2448	2	Q51338	Q51338 pseudomonas
421	7	1.5	950	4	Q9PZG0	Q9PZG0 homo sapien	494	7	1.5	2448	16	Q91182	Q91182 pseudomonas
422	7	1.5	959	16	Q9RLI2	Q9RLI2 streptomyce	495	7	1.5	3933	5	Q97239	Q97239 plasmodium
423	7	1.5	960	2	Q9RGV2	Q9RGV2 salmonella	496	7	1.5	4106	6	Q8WN21	Q8WN21 equus cabal
424	7	1.5	964	16	Q9AAH5	Q9AAH5 caulobacter	497	7	1.5	4144	6	Q8WN22	Q8WN22 canis famill
425	7	1.5	968	13	Q9PTI9	Q9PTI9 oncorhynch	498	7	1.5	4146	13	Q9DEI1	Q9DEI1 xenopus lae
426	7	1.5	976	12	Q9DUN0	Q9DUN0 kaposi's sa	499	7	1.5	4498	5	Q9M223	Q9M223 drosophila
427	7	1.5	978	5	Q9W542	Q9W542 drosophila	500	7	1.5	5183	4	Q8TDN5	Q8TDN5 homo sapien
428	7	1.5	979	13	Q93509	Q93509 xenopus lae	501	7	1.5	5183	2	Q9R5A6	Q9R5A6 thioabacillu
429	7	1.5	1003	12	Q91LX9	Q91LX9 kaposi's sa	502	7	1.5	502	22	Q9UCQ3	Q9UCQ3 drosophila
430	7	1.5	1022	5	Q9V7A9	Q9V7A9 drosophila	503	7	1.5	503	23	Q62594	Q62594 rattus leuc
431	7	1.5	1033	10	Q9XEX5	Q9XEX5 chlamydomon	504	7	1.5	504	24	Q9QVE1	Q9QVE1 rattus sp. cyt
432	7	1.5	1036	12	Q9DUM3	Q9DUM3 kaposi's sa	505	7	1.5	505	29	Q8XU13	Q8XU13 rattus sp. cyt
433	7	1.5	1062	5	Q9VTV2	Q9VTV2 drosophila	506	7	1.5	506	30	Q34903	Q34903 lastorinus
434	7	1.5	1070	3	Q9Y7B9	Q9Y7B9 kluyveromyc	507	7	1.5	507	31	Q4	Q4
435	7	1.5	1089	12	Q40947	Q40947 kaposi's sa	508	7	1.5	508	32	Q9UDE9	Q9UDE9 homo sapien
436	7	1.5	1112	13	Q9O381	Q9O381 carcharhinu	509	7	1.5	509	34	Q9KX7	Q9KX7 mus musculu
437	7	1.5	1114	4	Q9H0Y6	Q9H0Y6 homo sapien	510	7	1.5	510	34	Q9QV55	Q9QV55 mus sp. cyt
438	7	1.5	1129	12	Q9OR71	Q9OR71 kaposi's sa	511	7	1.5	511	36	Q9Q0S6	Q9Q0S6 tanapox vir
439	7	1.5	1134	4	Q9NPG3	Q9NPG3 homo sapien	512	7	1.5	512	37	Q9G1X8	Q9G1X8 mus musculu
440	7	1.5	1134	4	Q9E1P7	Q9E1P7 homo sapien	513	7	1.5	513	37	Q8UW71	Q8UW71 xenopus lae
441	7	1.5	1139	4	Q90280	Q90280 caenorhabd	514	7	1.5	514	42	Q9LE26	Q9LE26 psilocybe nu
442	7	1.5	1143	4	Q16614	Q16614 homo sapien	515	7	1.5	515	43	Q41997	Q41997 arabidopsis
443	7	1.5	1145	5	Q9VSP0	Q9VSP0 drosophila	516	7	1.5	516	44	P87665	P87665 duck adenov
444	7	1.5	1159	5	Q9VWC9	Q9VWC9 drosophila	517	7	1.5	517	45	Q49771	Q49771 arabidopsis
445	7	1.5	1160	2	Q9EUN8	Q9EUN8 corynebacte	518	7	1.5	518	46	Q8X3T3	Q8X3T3 escherichia
446	7	1.5	1162	12	Q98148	Q98148 kaposi's sa	519	7	1.5	519	47	Q942W8	Q942W8 caulobacter
447	7	1.5	1164	2	Q8VVJ4	Q8VVJ4 corynebacte	520	7	1.5	520	48	Q9VR85	Q9VR85 drosophila
448	7	1.5	1179	12	Q91L98	Q91L98 white spot	521	7	1.5	521	48	Q25451	Q25451 helicobacte
449	7	1.5	1180	12	Q8VAS9	Q8VAS9 white spot	522	7	1.5	522	48	Q92187	Q92187 canis famill
450	7	1.5	1185	4	Q8OTC5	Q8OTC5 white spot	523	7	1.5	523	49	Q28871	Q28871 canis famill
451	7	1.5	1204	4	Q9UHV4	Q9UHV4 homo sapien	524	7	1.5	524	49	Q98708	Q98708 rhizobium l
452	7	1.5	1217	11	Q9QW29	Q9QW29 rattus sp.	525	7	1.5	525	9	Q8SDM7	Q8SDM7 staphylococ
453	7	1.5	1218	11	Q99J66	Q99J66 rattus norv	526	7	1.5	526	16	Q99SQ2	Q99SQ2 staphylococ
454	7	1.5	1236	3	Q9C105	Q9C105 schizosacch	527	7	1.5	527	16	Q931J8	Q931J8 staphylococ

528	6	1.3	53	17	029381	029381 archaeoglob	601	6	1.3	80	5	Q9BP59	Q9BP59 conus ventr
529	6	1.3	54	12	099398	099398 hepatitis c	602	6	1.3	80	8	Q47577	Q47577 onchocerca
530	6	1.3	54	12	099404	099404 hepatitis c	603	6	1.3	80	10	Q08164	Q08164 secale cere
531	6	1.3	54	12	099403	099403 hepatitis c	604	6	1.3	80	12	Q70797	Q70797 tt virus . s
532	6	1.3	55	11	09ES47	09ES47 mus musculu	605	6	1.3	80	12	Q70803	Q70803 tt virus . s
533	6	1.3	55	16	08VK74	08VK74 mycobacteri	606	6	1.3	80	12	Q70809	Q70809 tt virus . s
534	6	1.3	56	2	Q9RFV2	Q9RFV2 salmonella	607	6	1.3	81	10	Q9M3G5	Q9M3G5 arabidopsis
535	6	1.3	57	16	Q8ZFW9	Q8ZFW9 yersinia pe	608	6	1.3	81	16	Q8XLM8	Q8XLM8 clostridium
536	6	1.3	57	17	Q974H3	Q974H3 sulfolobus	609	6	1.3	83	7	Q19220	Q19220 rangifer ta
537	6	1.3	58	2	Q47832	Q47832 enterococcu	610	6	1.3	83	7	Q19221	Q19221 rangifer ta
538	6	1.3	58	11	Q9CR05	Q9CR05 mus musculu	611	6	1.3	83	7	Q19222	Q19222 rangifer ta
539	6	1.3	58	11	Q9JH07	Q9JH07 mus musculu	612	6	1.3	83	7	Q19227	Q19227 rangifer ta
540	6	1.3	59	4	Q43705	Q43705 homo sapien	613	6	1.3	83	9	Q38065	Q38065 bacterioph
541	6	1.3	59	6	Q28675	Q28675 oryctolagus	614	6	1.3	83	16	Q915K6	Q915K6 pseudomonas
542	6	1.3	61	2	P94471	P94471 bacillus su	615	6	1.3	83	16	Q8RF71	Q8RF71 bruceella me
543	6	1.3	61	5	Q9NFP9	Q9NFP9 ceratitlis c	616	6	1.3	85	4	Q9U0B0	Q9U0B0 homo sapien
544	6	1.3	61	5	Q9NFP8	Q9NFP8 ceratitlis c	617	6	1.3	85	4	Q95419	Q95419 homo sapien
545	6	1.3	61	5	Q8STG3	Q8STG3 echinococcu	618	6	1.3	85	5	Q44629	Q44629 caenorhabdi
546	6	1.3	61	11	Q9ES48	Q9ES48 mus musculu	619	6	1.3	85	5	Q8T0W5	Q8T0W5 plasmia hypo
547	6	1.3	62	13	Q91801	Q91801 xenopus lae	620	6	1.3	85	10	Q8VZW8	Q8VZW8 arabidopsis
548	6	1.3	63	5	Q9VBH1	Q9VBH1 drosophila	621	6	1.3	85	12	Q98433	Q98433 parametium
549	6	1.3	65	15	Q12385	Q12385 simian t-ly	622	6	1.3	85	12	Q99GX8	Q99GX8 helicoverpa
550	6	1.3	66	17	Q8TW36	Q8TW36 methanopyru	623	6	1.3	86	2	Q9S4W5	Q9S4W5 escherichia
551	6	1.3	67	2	Q87442	Q87442 pseudomonas	624	6	1.3	86	3	Q9C2X8	Q9C2X8 candida alb
552	6	1.3	67	3	Q94112	Q94112 pneumocysti	625	6	1.3	86	15	Q8USX0	Q8USX0 human immun
553	6	1.3	67	5	Q02400	Q02400 teladorsagi	626	6	1.3	86	16	Q8UXA0	Q8UXA0 agrobacteri
554	6	1.3	68	8	Q9MOK8	Q9MOK8 hydra atten	627	6	1.3	87	15	Q80794	Q80794 human t-lym
555	6	1.3	68	17	Q9HM00	Q9HM00 thermoplasma	628	6	1.3	87	15	Q9PYZ8	Q9PYZ8 human t-lym
556	6	1.3	69	12	Q67840	Q67840 hepatitis b	629	6	1.3	87	15	Q9PXY6	Q9PXY6 human t-lym
557	6	1.3	69	12	Q67846	Q67846 hepatitis b	630	6	1.3	87	16	Q92TX1	Q92TX1 rhizobium m
558	6	1.3	69	12	Q67847	Q67847 hepatitis b	631	6	1.3	88	15	Q9OBK4	Q9OBK4 human immun
559	6	1.3	69	12	Q67848	Q67848 hepatitis b	632	6	1.3	89	2	Q9ATF6	Q9ATF6 actinobactil
560	6	1.3	69	13	Q8OR05	Q8OR05 rana pipien	633	6	1.3	89	10	Q8S6T9	Q8S6T9 oryza sativ
561	6	1.3	69	16	Q9A482	Q9A482 caulobacter	634	6	1.3	89	16	Q9USK2	Q9USK2 chlamydia p
562	6	1.3	69	16	Q8RD68	Q8RD68 thermotanaer	635	6	1.3	89	16	Q9CE15	Q9CE15 lactococcus
563	6	1.3	71	12	Q93PY4	Q93PY4 pseudomonas	636	6	1.3	90	16	P72713	P72713 synecocyst
564	6	1.3	71	12	Q8VAS3	Q8VAS3 white spot	637	6	1.3	90	16	Q8YBM5	Q8YBM5 bruceella me
565	6	1.3	71	12	Q8ORC3	Q8ORC3 white spot	638	6	1.3	91	16	Q8SBR5	Q8SBR5 arabidopsis
566	6	1.3	71	13	Q8ORC4	Q8ORC4 rana pipien	639	6	1.3	91	16	Q9CCB9	Q9CCB9 mycobacteri
567	6	1.3	71	17	Q971H4	Q971H4 sulfolobus	640	6	1.3	91	16	Q8X7S2	Q8X7S2 escherichia
568	6	1.3	72	3	Q9C174	Q9C174 aspergillus	641	6	1.3	92	10	Q9A6S5	Q9A6S5 oryza sativ
569	6	1.3	72	15	P89228	P89228 walleye der	642	6	1.3	93	16	Q934X2	Q934X2 salmonella
570	6	1.3	72	16	Q97GV7	Q97GV7 clostridium	643	6	1.3	94	4	Q14057	Q14057 homo sapien
571	6	1.3	72	16	Q8ZJ13	Q8ZJ13 yersinia pe	644	6	1.3	94	5	Q9U705	Q9U705 plasmidium
572	6	1.3	74	5	Q25016	Q25016 helix asper	645	6	1.3	94	5	Q8U708	Q8U708 plasmidium
573	6	1.3	74	10	Q8S221	Q8S221 oryza sativ	646	6	1.3	94	5	Q9U706	Q9U706 plasmidium
574	6	1.3	74	17	Q974F5	Q974F5 sulfolobus	647	6	1.3	94	5	Q9U705	Q9U705 plasmidium
575	6	1.3	75	5	Q9BP58	Q9BP58 conus ventr	648	6	1.3	94	17	Q8TK50	Q8TK50 methanosarc
576	6	1.3	75	12	Q84587	Q84587 parametium	649	6	1.3	96	5	Q9VK38	Q9VK38 drosophila
577	6	1.3	76	2	Q9FB50	Q9FB50 corynebacte	650	6	1.3	96	11	Q8R444	Q8R444 phodopus su
578	6	1.3	76	5	Q8GPN6	Q8GPN6 schistosoma	651	6	1.3	96	12	Q98554	Q98554 parametium
579	6	1.3	76	5	Q9BP60	Q9BP60 conus ventr	652	6	1.3	96	17	Q8TL15	Q8TL15 methanosarc
580	6	1.3	76	10	Q9ASL0	Q9ASL0 oryza sativ	653	6	1.3	98	8	Q950L7	Q950L7 rhizophydlu
581	6	1.3	76	12	Q98541	Q98541 parametium	654	6	1.3	98	15	Q9END6	Q9END6 human immun
582	6	1.3	76	16	Q9JTT8	Q9JTT8 neisseria m	655	6	1.3	98	16	P73230	P73230 synecocyst
583	6	1.3	77	5	Q9BP68	Q9BP68 conus arena	656	6	1.3	98	16	Q92CK3	Q92CK3 listeria in
584	6	1.3	77	5	Q9BP65	Q9BP65 conus arena	657	6	1.3	98	16	Q8RT86	Q8RT86 listeria mo
585	6	1.3	77	16	Q816P4	Q816P4 listeria mo	658	6	1.3	99	12	Q91NV0	Q91NV0 hepatitis b
586	6	1.3	78	5	Q9BP76	Q9BP76 conus tessu	659	6	1.3	99	12	Q91NV8	Q91NV8 hepatitis b
587	6	1.3	78	5	Q9BP66	Q9BP66 conus arena	660	6	1.3	99	12	Q91NWO	Q91NWO hepatitis b
588	6	1.3	78	5	Q9BP64	Q9BP64 conus penna	661	6	1.3	99	12	Q91NV8	Q91NV8 hepatitis b
589	6	1.3	78	7	Q8WM11	Q8WM11 rangifer ta	662	6	1.3	99	12	Q91NV6	Q91NV6 hepatitis b
590	6	1.3	78	9	Q8WM08	Q8WM08 rangifer ta	663	6	1.3	99	12	Q91NV4	Q91NV4 hepatitis b
591	6	1.3	78	9	Q9MB09	Q9MB09 staphylococ	664	6	1.3	99	12	Q73557	Q73557 lassia virus
592	6	1.3	78	10	Q9M5S9	Q9M5S9 ipomoea bat	665	6	1.3	100	2	Q9RPF5	Q9RPF5 streptomyce
593	6	1.3	79	5	Q9BP73	Q9BP73 conus tessu	666	6	1.3	100	2	Q93TB3	Q93TB3 rhodobacter
594	6	1.3	79	5	Q9BP72	Q9BP72 conus arena	667	6	1.3	100	2	Q8RPY7	Q8RPY7 rhizobium l
595	6	1.3	79	5	Q9BP71	Q9BP71 conus arena	668	6	1.3	100	2	Q8RPY7	Q8RPY7 rhizobium l
596	6	1.3	79	5	Q9BP70	Q9BP70 conus arena	669	6	1.3	100	8	Q96640	Q96640 ceratitlis c
597	6	1.3	79	5	Q9BP69	Q9BP69 conus arena	670	6	1.3	100	8	Q8WAN2	Q8WAN2 hemisplinus
598	6	1.3	79	5	Q9BP67	Q9BP67 conus arena	671	6	1.3	100	15	Q80322	Q80322 human immun
599	6	1.3	79	5	Q9BP63	Q9BP63 conus penna	672	6	1.3	100	16	Q98CY3	Q98CY3 rhizobium l
600	6	1.3	80	5	Q9BP62	Q9BP62 conus ventr	673	6	1.3	100	17	Q28004	Q28004 archaeoglob

674	6	1.3	101	8	020826	020826	Ierista dor	747	6	1.3	114	15	09YL60	09YL60 human immun
675	6	1.3	101	16	092GB3	092GB3	rickettsia	748	6	1.3	114	17	096ZM3	096ZM3 sulfolobus
676	6	1.3	102	10	09SBH8	09SBH8	solanum tub	749	6	1.3	115	14	08VMW7	08VMW7 listeria se
677	6	1.3	102	15	077143	077143	human immun	750	6	1.3	115	4	09PIU3	09PIU3 homo sapien
678	6	1.3	102	17	096XB9	096XB9	sulfolobus	751	6	1.3	115	4	09S325	09S325 homo sapien
679	6	1.3	103	2	093EX2	093EX2	rhodococcus	752	6	1.3	115	10	09S9A1	09S9A1 viscum albu
680	6	1.3	103	2	049853	049853	mycobacteri	753	6	1.3	115	15	09WLS9	09WLS9 human immun
681	6	1.3	103	5	09U3J9	09U3J9	caenorhabdi	754	6	1.3	115	15	09YL38	09YL38 human immun
682	6	1.3	103	5	09VY03	09VY03	drosophila	755	6	1.3	115	15	080318	080318 human immun
683	6	1.3	103	10	09M045	09M045	arabidopsis	756	6	1.3	115	15	070106	070106 human immun
684	6	1.3	103	16	09PH85	09PH85	xyliella fas	757	6	1.3	115	15	P88778	P88778 human immun
685	6	1.3	103	16	006257	006257	mycobacteri	758	6	1.3	115	16	08YTX8	08YTX8 anaebaena sp
686	6	1.3	104	10	09PEX6	09PEX6	hordium vul	759	6	1.3	115	17	09YET6	09YET6 aetopyrum p
687	6	1.3	104	16	033291	033291	pyrococcul	760	6	1.3	115	17	097AU6	097AU6 thermoplas
688	6	1.3	104	17	08U0B2	08U0B2	cornebacte	761	6	1.3	116	6	09BCS3	09BCS3 macaca fasc
689	6	1.3	105	2	08RLM2	08RLM2	caenorhabdi	762	6	1.3	116	8	037127	037127 arabcia ltx
690	6	1.3	105	5	08W0E9	08W0E9	egyptian su	763	6	1.3	116	17	09Y8P4	09Y8P4 aetopyrum p
691	6	1.3	105	12	056310	056310	egyptian su	764	6	1.3	117	8	009315	009315 eidmannella
692	6	1.3	105	12	09YX2	09YX2	egyptian su	765	6	1.3	117	10	09LH43	09LH43 arabidopsis
693	6	1.3	105	12	09YXD8	09YXD8	egyptian su	766	6	1.3	117	12	08QP43	08QP43 dengue viru
694	6	1.3	105	12	09YXD5	09YXD5	egyptian su	767	6	1.3	118	10	09LQ03	09LQ03 arabidopsis
695	6	1.3	105	12	09YXD5	09YXD5	egyptian su	768	6	1.3	119	2	093RT7	093RT7 thermotoga
696	6	1.3	105	16	09Z9D1	09Z9D1	chlamydia p	769	6	1.3	119	2	093RT5	093RT5 thermotoga
697	6	1.3	105	16	08Y2P7	08Y2P7	raistonia s	770	6	1.3	119	3	093RD4	093RD4 thermotoga
698	6	1.3	106	8	P92346	P92346	arabidopsis	771	6	1.3	119	3	093RX7	093RX7 schizosacch
699	6	1.3	106	10	P93Z99	P93Z99	arabidopsis	772	6	1.3	119	5	08T5W1	08T5W1 caenorhabdi
700	6	1.3	106	16	08ZBP5	08ZBP5	yersinia pe	773	6	1.3	119	5	08T5V5	08T5V5 caenorhabdi
701	6	1.3	106	17	08TML6	08TML6	methanosarc	774	6	1.3	119	5	08T5V4	08T5V4 caenorhabdi
702	6	1.3	107	16	08UCS6	08UCS6	agrobacteri	775	6	1.3	119	12	09YH78	09YH78 cherry necr
703	6	1.3	108	3	013531	013531	saccharomyc	776	6	1.3	119	15	09U678	09U678 human immun
704	6	1.3	108	12	084Z59	084Z59	egyptian su	777	6	1.3	120	2	032318	032318 bacillus th
705	6	1.3	108	16	091097	091097	pseudomonas	778	6	1.3	120	5	08T5W2	08T5W2 caenorhabdi
706	6	1.3	108	16	08Z7M2	08Z7M2	salmonella	779	6	1.3	120	8	099150	099150 rivulus del
707	6	1.3	108	17	09YB07	09YB07	aetopyrum p	780	6	1.3	120	8	099151	099151 rivulus sp.
708	6	1.3	109	2	P71157	P71157	desulfofibr	781	6	1.3	120	8	099157	099157 rivulus apl
709	6	1.3	109	2	043891	043891	acetobacter	782	6	1.3	120	8	099158	099158 rivulus pun
710	6	1.3	109	3	013543	013543	saccharomyc	783	6	1.3	120	8	099164	099164 neofundulus
711	6	1.3	109	12	09WAZ5	09WAZ5	lt virus. o	784	6	1.3	120	8	035732	035732 rivulus mag
712	6	1.3	109	16	09PEF4	09PEF4	xyliella fas	785	6	1.3	120	8	035711	035711 rivulus cry
713	6	1.3	109	17	08TR36	08TR36	methanosarc	786	6	1.3	120	8	P87398	P87398 tritonectes
714	6	1.3	110	11	09D2G6	09D2G6	mus musculu	787	6	1.3	120	8	099142	099142 rivulus fire
715	6	1.3	110	16	09CDP7	09CDP7	lactococcus	788	6	1.3	120	12	067428	067428 dengue viru
716	6	1.3	110	17	09Y1P0	09Y1P0	pyrococcus	789	6	1.3	120	16	098PB3	098PB3 rhizobium l
717	6	1.3	111	4	014502	014502	homo sapien	790	6	1.3	120	16	053437	053437 mycobacteri
718	6	1.3	111	12	09WAZ3	09WAZ3	lt virus. o	791	6	1.3	121	4	08TCU2	08TCU2 homo sapien
719	6	1.3	111	12	09WAZ7	09WAZ7	lt virus. o	792	6	1.3	121	8	09TGS4	09TGS4 rhinolophus
720	6	1.3	111	12	08OSA0	08OSA0	felis domes	793	6	1.3	121	8	09BPD6	09BPD6 cilix glauc
721	6	1.3	111	16	005219	005219	baetilia su	794	6	1.3	121	16	08XBE4	08XBE4 escherichia
722	6	1.3	111	16	08T0B4	08T0B4	raistonia s	795	6	1.3	121	16	08RI03	08RI03 fusobacteri
723	6	1.3	112	8	0922L3	0922L3	uta antiqua	796	6	1.3	122	2	09P928	09P928 pseudomonas
724	6	1.3	112	9	080205	080205	methanobact	797	6	1.3	122	2	049696	049696 mycobacteri
725	6	1.3	112	10	09FEW6	09FEW6	oryza sativ	798	6	1.3	122	10	P82178	P82178 spinacia ol
726	6	1.3	112	10	08RZU6	08RZU6	oryza sativ	799	6	1.3	122	16	067205	067205 aquilex aeo
727	6	1.3	112	11	060880	060880	mus musculu	800	6	1.3	122	16	09RTW6	09RTW6 deinococcus
728	6	1.3	112	11	08R8E9	08R8E9	mus musculu	801	6	1.3	123	2	09RC10	09RC10 salmonella
729	6	1.3	112	16	08ZSA9	08ZSA9	anaebaena sp	802	6	1.3	123	11	09CSU8	09CSU8 mus musculu
730	6	1.3	112	17	09YDD6	09YDD6	aetopyrum p	803	6	1.3	124	2	09RLT0	09RLT0 lactobacill
731	6	1.3	113	2	045610	045610	baetilia su	804	6	1.3	124	16	08XYI4	08XYI4 clostridium
732	6	1.3	113	13	012967	012967	fugu rubrip	805	6	1.3	124	17	057846	057846 pyrococcus
733	6	1.3	113	15	070647	070647	simian t-ly	806	6	1.3	125	5	045212	045212 brugia paha
734	6	1.3	113	16	09K9L4	09K9L4	baetilia ha	807	6	1.3	125	10	09SWM1	09SWM1 argyroxiphil
735	6	1.3	114	3	09HEN7	09HEN7	cocciobolu	808	6	1.3	125	12	09DM60	09DM60 rat cytomeg
736	6	1.3	114	3	09HEN5	09HEN5	cocciobolu	809	6	1.3	125	17	059155	059155 pyrococcus
737	6	1.3	114	3	09HEP2	09HEP2	bipolaris s	810	6	1.3	126	17	08TWM0	08TWM0 methanosarc
738	6	1.3	114	3	003502	003502	kluyveromyc	811	6	1.3	127	4	09HB18	09HB18 homo sapien
739	6	1.3	114	5	08T5V8	08T5V8	caenorhabdi	812	6	1.3	127	8	09PE60	09PE60 homo sapien
740	6	1.3	114	8	09XP27	09XP27	suaeda japo	813	6	1.3	127	10	09MP37	09MP37 ctenomys bo
741	6	1.3	114	10	09SDS3	09SDS3	capsicum an	814	6	1.3	127	16	098390	098390 oryza sativ
742	6	1.3	114	10	09SDS2	09SDS2	lycopersico	815	6	1.3	127	16	092RT8	092RT8 rhizobium m
743	6	1.3	114	10	024037	024037	lycopersico	816	6	1.3	127	16	09AJZ0	09AJZ0 streptomyc
744	6	1.3	114	10	024038	024038	lycopersico	817	6	1.3	128	1	033155	033155 methanosarc
745	6	1.3	114	12	08VTE6	08VTE6	lt virus. o	818	6	1.3	128	6	09MYS8	09MYS8 oryctolagus
746	6	1.3	114	12	08UYC2	08UYC2	lt virus. o	819	6	1.3	128	8	P92832	P92832 colastes ln



820	6	1.3	128	11	Q9D1X3	Q9d1x3 mus musculu	893	6	1.3	139	16	Q9RSF2	Q9rsf2 deinooccus
821	6	1.3	128	12	Q8V7E4	Q8v7e4 tt virus. o	894	6	1.3	139	16	Q9K1V0	Q9k1v0 vibrio chol
822	6	1.3	128	12	Q8V7C7	Q8v7c7 tt virus. o	895	6	1.3	140	10	Q9SHZ7	Q9shz7 arabidopsis
823	6	1.3	128	12	Q8UY10	Q8uy10 tt virus. o	896	6	1.3	140	11	Q9D1M5	Q9d1m5 mus musculu
824	6	1.3	128	16	Q9PMX6	Q9pmx6 campylobact	897	6	1.3	140	11	Q8XV77	Q8xv77 ralsstonia s
825	6	1.3	129	8	Q9B0N3	Q9b0n3 paragonimus	898	6	1.3	141	9	Q9MCJ2	Q9mcj2 streptococc
826	6	1.3	129	8	Q03367	Q03367 paragonimus	899	6	1.3	141	16	Q60148	Q60148 pseudomonas
827	6	1.3	129	8	Q9B3H5	Q9b3h5 paragonimus	900	6	1.3	141	16	Q9CFA6	Q9cfa6 lactococcus
828	6	1.3	129	10	Q22933	Q22933 arabidopsis	901	6	1.3	141	16	Q8XC15	Q8xc15 escherichia
829	6	1.3	129	10	Q9FWY6	Q9fwy6 arabidopsis	902	6	1.3	142	10	Q9FS06	Q9fs06 triticum ae
830	6	1.3	129	11	Q9ESG1	Q9esg1 mus musculu	903	6	1.3	142	10	Q9LF87	Q9lf87 arabidopsis
831	6	1.3	129	15	Q36916	Q36916 human immun	904	6	1.3	142	11	Q99NK6	Q99nk6 cricetus
832	6	1.3	129	15	Q9Q8K7	Q9q8k7 human immun	905	6	1.3	142	13	Q9PWP2	Q9pwp2 xenopus lae
833	6	1.3	129	15	Q9Q8K6	Q9q8k6 human immun	906	6	1.3	142	13	Q9PWP1	Q9pwp1 xenopus lae
834	6	1.3	129	15	Q9Q8K5	Q9q8k5 human immun	907	6	1.3	142	13	Q9PWP1	Q9pwp1 xenopus lae
835	6	1.3	129	15	Q9Q8K5	Q9q8k5 human immun	907	6	1.3	142	13	Q9PWP1	Q9pwp1 xenopus lae
836	6	1.3	129	17	Q8U237	Q8u237 pyrococcus	908	6	1.3	142	16	Q9Z0R0	Q9z0r0 rhizobium m
837	6	1.3	130	2	Q9R918	Q9r918 bacillus su	909	6	1.3	142	17	Q9YEO9	Q9yeo9 aeropyrum p
838	6	1.3	130	2	Q9L7W8	Q9l7w8 bacillus su	910	6	1.3	142	17	Q9YAO0	Q9yao0 aeropyrum p
839	6	1.3	130	8	Q03358	Q03358 paragonimus	911	6	1.3	142	17	Q9HMA4	Q9hma4 halobacteri
840	6	1.3	130	8	Q03358	Q03358 paragonimus	911	6	1.3	142	17	Q9HMA4	Q9hma4 halobacteri
841	6	1.3	130	8	Q03358	Q03358 paragonimus	912	6	1.3	143	5	Q20396	Q20396 caenorhabdi
842	6	1.3	130	8	Q03358	Q03358 paragonimus	912	6	1.3	143	5	Q20396	Q20396 caenorhabdi
843	6	1.3	130	8	Q03358	Q03358 paragonimus	912	6	1.3	143	5	Q20396	Q20396 caenorhabdi
844	6	1.3	130	8	Q21214	Q21214 paragonimus	913	6	1.3	143	5	Q20396	Q20396 caenorhabdi
845	6	1.3	130	8	Q21225	Q21225 paragonimus	913	6	1.3	143	5	Q20396	Q20396 caenorhabdi
846	6	1.3	130	8	Q9TEX4	Q9tex4 paragonimus	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
847	6	1.3	130	8	Q9TEX3	Q9tex3 paragonimus	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
848	6	1.3	130	8	Q9TEX2	Q9tex2 paragonimus	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
849	6	1.3	130	8	Q9TEX1	Q9tex1 paragonimus	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
850	6	1.3	130	8	Q9TEX0	Q9tex0 paragonimus	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
851	6	1.3	130	8	Q03368	Q03368 paragonimus	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
852	6	1.3	130	8	Q03365	Q03365 paragonimus	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
853	6	1.3	130	10	Q9M3C3	Q9m3c3 arabidopsis	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
854	6	1.3	130	16	Q34506	Q34506 bacillus su	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
855	6	1.3	130	17	Q29584	Q29584 archaeoglob	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
856	6	1.3	131	2	Q50281	Q50281 pseudomonas	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
857	6	1.3	131	2	Q93137	Q93137 streptomyc	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
858	6	1.3	131	5	Q960H9	Q960h9 drosophila	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
859	6	1.3	131	8	Q35865	Q35865 strongyloce	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
860	6	1.3	131	8	Q35903	Q35903 strongyloce	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
861	6	1.3	131	10	Q9FRL0	Q9frl0 arabidopsis	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
862	6	1.3	131	11	Q9CZL2	Q9czl2 mus musculu	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
863	6	1.3	131	16	Q8XZ4	Q8xz4 salmonella	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
864	6	1.3	132	4	Q8TBF9	Q8tbf9 homo sapien	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
865	6	1.3	132	16	Q928H3	Q928h3 listeria in	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
866	6	1.3	132	16	Q8YT00	Q8yt00 anabaena sp	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
867	6	1.3	133	5	Q8T668	Q8t668 echinococcu	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
868	6	1.3	133	5	Q8T667	Q8t667 echinococcu	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
869	6	1.3	133	8	Q35820	Q35820 strongyloce	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
870	6	1.3	133	15	Q9Q0R7	Q9q0r7 human immun	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
871	6	1.3	133	17	Q59150	Q59150 pyrococcus	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
872	6	1.3	134	2	Q91895	Q91895 prochloroco	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
873	6	1.3	134	12	Q9DUH9	Q9duh9 tt virus. o	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
874	6	1.3	134	13	Q98S16	Q98s16 brachydanio	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
875	6	1.3	134	16	Q9KPO2	Q9kpo2 vibrio chol	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
876	6	1.3	134	16	Q98EO3	Q98eo3 rhizobium i	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
877	6	1.3	135	5	Q8SUW0	Q8suw0 encephalito	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
878	6	1.3	135	16	P72881	P72881 synchocyst	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
879	6	1.3	135	16	Q9A142	Q9a142 streptococc	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
880	6	1.3	135	17	Q9YCX8	Q9ycx8 aeropyrum p	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
881	6	1.3	136	4	Q9Y423	Q9y423 homo sapien	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
882	6	1.3	136	10	Q9FWZ6	Q9fwz6 oryza sativ	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
883	6	1.3	136	10	Q946W7	Q946w7 prunus pers	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
884	6	1.3	136	11	Q9CZJ4	Q9czj4 mus musculu	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
885	6	1.3	136	16	Q97NT9	Q97nt9 streptococc	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
886	6	1.3	136	16	Q8U8L2	Q8u8l2 agrobacteri	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
887	6	1.3	137	17	Q58445	Q58445 pyrococcus	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
888	6	1.3	138	16	Q930H6	Q930h6 rhizobium m	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
889	6	1.3	138	16	Q9AD53	Q9ad53 streptomyc	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
890	6	1.3	138	17	Q9YC20	Q9yc20 aeropyrum p	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
891	6	1.3	138	17	Q8U224	Q8u224 pyrococcus	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi
892	6	1.3	139	5	Q9XX29	Q9xx29 caenorhabdi	914	6	1.3	143	5	Q20396	Q20396 caenorhabdi

```

966 6 1.3 152 4 09BxF8 09BxF8 homo sapien
967 6 1.3 152 12 09DUD0 09DUD0 tt virus. o
968 6 1.3 152 16 09KDA2 09KDA2 bacillus. o
969 6 1.3 152 16 08UJ34 08UJ34 agrobacteri
970 6 1.3 153 8 09T5G5 09T5G5 teladorsagi
971 6 1.3 153 9 09J184 09J184 parapatilase
972 6 1.3 153 9 09XJ95 09XJ95 streptococ
973 6 1.3 153 10 09LGM4 09LGM4 oryza sativ
974 6 1.3 153 12 09DS06 09DS06 foot-and-mo
975 6 1.3 153 12 08V5P5 08V5P5 foot-and-mo
976 6 1.3 153 12 08V5P4 08V5P4 foot-and-mo
977 6 1.3 153 12 08V5N7 08V5N7 foot-and-mo
978 6 1.3 153 12 08V5N6 08V5N6 foot-and-mo
979 6 1.3 153 12 08V0P7 08V0P7 foot-and-mo
980 6 1.3 153 16 09X0B4 09X0B4 thermotoga
981 6 1.3 153 16 09JR28 09JR28 neisseria m
982 6 1.3 153 16 08VP90 08VP90 anabaena sp
983 6 1.3 153 17 0978F8 0978F8 thermoplas
984 6 1.3 154 4 09H4G4 09H4G4 homo sapien
985 6 1.3 154 8 09G6Z6 09G6Z6 haplopharyn
986 6 1.3 154 11 09CYL5 09CYL5 mus musculu
987 6 1.3 154 13 09DFE0 09DFE0 oncorhynchu
988 6 1.3 154 16 09KO58 09KO58 vibrio chol
989 6 1.3 154 16 08ZKE6 08ZKE6 salmonella
990 6 1.3 154 16 08Z1P8 08Z1P8 salmonella
991 6 1.3 155 8 0479S1 0479S1 oryza sativ
992 6 1.3 155 12 099AR2 099AR2 tt virus. o
993 6 1.3 156 11 06Z571 06Z571 canine herp
994 6 1.3 156 12 066091 066091 tt virus. o
995 6 1.3 156 12 099AR6 099AR6 tt virus. o
996 6 1.3 157 6 08WN54 08WN54 sus scrofa
997 6 1.3 157 8 09X1H7 09X1H7 rhinolophus
998 6 1.3 157 8 09X1H9 09X1H9 rhinolophus
999 6 1.3 157 10 08S3Z1 08S3Z1 sorghum b1c
1000 6 1.3 157 12 08UZK8 08UZK8 sen virus.

```

## ALIGNMENTS

```

RESULT 1
09UIFO PRELIMINARY; PRT; 289 AA.
ID 09UIFO
AC 09UIFO
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Presentin 1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99063792; PubMed=9847074;
RA Sulston J.E.; Waterston R.;
RT "Toward a complete human genome sequence.";
RL Genome Res. 8:1097-1108(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX Kozlowicz A.; Pape K.; Biewald T.;
RT "The sequence of Homo sapiens PAC clone RP4-687K1.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004858; AF19253.1; -.
DR InterPro; IPR001108; Presentin.
DR Pfam; PFO1080; Presentin.1.
DR PRINTS; PRO1072; PRESENTIN.
FT NON_TER 289
SQ SEQUENCE 289 AA; 33067 MW; D503395A31C0ECB CRC64;

```

```

Query Match 43.7%; Score 204; DB 4; Length 289;
Best Local Similarity 100.0%; Pred. No. 2e-199;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MTELPAPLSTFONAMQMSDNLSTNVRSONDRERQEHNDRRSLGHPPLNSGRPOGNSR 60
DB 1 MTELPAPLSTFONAMQMSDNLSTNVRSONDRERQEHNDRRSLGHPPLNSGRPOGNSR 60
OY 61 QVEDDEDEDELTKYAKAHVIMLFVPVTLQMVVATIKSVSPYTRKDDQLYTPFTE 120
DB 61 QVEDDEDEDELTKYAKAHVIMLFVPVTLQMVVATIKSVSPYTRKDDQLYTPFTE 120
OY 121 DTEVVGQALSHLSILNAALMISVIVMTLLVLYKRCYKVIHAWLIITSLLLFFFSFI 180
DB 121 DTEVVGQALSHLSILNAALMISVIVMTLLVLYKRCYKVIHAWLIITSLLLFFFSFI 180
OY 181 YLGEVFKTYNVAVDYITVALLIWN 204
DB 181 YLGEVFKTYNVAVDYITVALLIWN 204

```

## RESULT 2

```

ID 095465 PRELIMINARY; PRT; 184 AA.
AC 095465;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE MINILIN.
GN PSNI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX Powell C.S.; Gegg M.E.; Palmer M.S.;
RT "Human presentin 1 gene encodes an alternative protein-minilin.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99063792; PubMed=9847074;
RA Sulston J.E.; Waterston R.;
RT "Toward a complete human genome sequence.";
RL Genome Res. 8:1097-1108(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX Kozlowicz A.; Pape K.; Biewald T.;
RT "The sequence of Homo sapiens PAC clone RP4-687K1.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Waterston R.H.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AU008005; CA07825.1; -.
DR EMBL; AC004858; AF19254.1; -.
DR InterPro; IPR001108; Presentin.
DR Pfam; PFO1080; Presentin.1.
SQ SEQUENCE 184 AA; 21073 MW; 95C68A7EA0020874 CRC64;

```

```

Query Match 34.5%; Score 161; DB 4; Length 184;
Best Local Similarity 100.0%; Pred. No. 1e-155;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 MTELPAPLSTFONAMQMSDNLSTNVRSONDRERQEHNDRRSLGHPPLNSGRPOGNSR 60
DB 1 MTELPAPLSTFONAMQMSDNLSTNVRSONDRERQEHNDRRSLGHPPLNSGRPOGNSR 60

```

Qy 61 QVVEDEDEDELTKYGAHVIMLFVPTLCMVVVAATIKSVSFYTRKDGQIYTPETE 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 61 QVVEDEDEDELTKYGAHVIMLFVPTLCMVVVAATIKSVSFYTRKDGQIYTPETE 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 121 DTEVVGORALHSILMAIMISIVYVMTLLVYLYKRYCKV 161  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 121 DTEVVGORALHSILMAIMISIVYVMTLLVYLYKRYCKV 161  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

## RESULT 3

Q96P33 PRELIMINARY: PRT: 378 AA.  
 AC Q96P33;  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 DE Presentin 1.  
 GN PSEN1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NC NCBL\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kang L., Zhang B., Zhou Y., Peng X., Yuan J., Qiang B.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF416717; AAL16811.1; -  
 DR InterPro: IPR001108; Presentin.  
 DR Pfam: PF01080; Presentin.1.  
 DR PRINTS: PRO1072; PRESENTIN.  
 SQ SEQUENCE 378 AA; 42658 MW; A09D6D2FC6F5618D CRC64;

Query Match 27.2%; Score 127; DB 4; Length 378;  
 Best Local Similarity 100.0%; Pred. No. 9, 8e-121;  
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTELPALSYFQNAQMSQEDNHLSTNVSQNDNREROEHNDRLSGHPEPLSNGRPGNSR 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1 MTELPALSYFQNAQMSQEDNHLSTNVSQNDNREROEHNDRLSGHPEPLSNGRPGNSR 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 61 QVVEDEDEDELTKYGAHVIMLFVPTLCMVVVAATIKSVSFYTRKDGQIYTPETE 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 61 QVVEDEDEDELTKYGAHVIMLFVPTLCMVVVAATIKSVSFYTRKDGQIYTPETE 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 121 DTEVVG 127  
 |||||||  
 Db 121 DTEVVG 127

## RESULT 4

Q91WK6 PRELIMINARY: PRT: 261 AA.  
 AC Q91WK6;  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Similar to presentin 1.  
 GN PSEN1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NC NCBL\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Tissue-eye, AND RETINA;  
 RA Strausberg R.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC014744; AAL14744.1; -  
 DR MGD: MGI:1202717; Psen1.  
 DR InterPro: IPR001108; Presentin.  
 DR Pfam: PF01080; Presentin.1.  
 SQ SEQUENCE 261 AA; 29849 MW; 16FB53FBE20EA4EB CRC64;

Query Match 18.6%; Score 87; DB 11; Length 261;  
 Best Local Similarity 100.0%; Pred. No. 4, 7e-80;  
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 NSRQVDEDEDELTKYGAHVIMLFVPTLCMVVVAATIKSVSFYTRKDGQIYTP 117  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 58 NSRQVDEDEDELTKYGAHVIMLFVPTLCMVVVAATIKSVSFYTRKDGQIYTP 117  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 118 FTEDTEVVGORALHSILMAIMISIV 144  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 118 FTEDTEVVGORALHSILMAIMISIV 144  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

## RESULT 5

Q90X08 PRELIMINARY: PRT: 468 AA.  
 AC Q90X08;  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 DE Presentin 1.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 NC NCBL\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EL4; TISSUE=BRAIN;  
 RA Korade Mirnics Z., Keryanov S., Lovelock J., Corey S.J.;  
 RT "Cloning of chicken presentin.  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY043492; AAK95408.1; -  
 DR InterPro: IPR001108; Presentin.  
 DR Pfam: PF01080; Presentin.1.  
 DR PRINTS: PRO1072; PRESENTIN.  
 SQ SEQUENCE 468 AA; 52812 MW; B746BDA3BAC0BFA CRC64;

Query Match 13.5%; Score 63; DB 13; Length 468;  
 Best Local Similarity 100.0%; Pred. No. 2, 4e-55;  
 Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 213 IHMKGPLRLOQAVYIMISALMALVFIKYPMTAMLIYAVSYDVAVLCPPKGPLRLV 272  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 214 IHMKGPLRLOQAVYIMISALMALVFIKYPMTAMLIYAVSYDVAVLCPPKGPLRLV 273  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 273 ETA 275  
 ||||  
 Db 274 ETA 276

## RESULT 6

Q73869 PRELIMINARY: PRT: 384 AA.  
 AC Q73869;  
 DT 01-AUG-1998 (TReMBLrel. 07, Created)  
 DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE Psi protein (Fragment).  
 GN Psi.  
 OS Cyprinus carpio (Common carp).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Cyprinus.  
 NC NCBL\_TaxID=7962;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98337349; PubMed=9674592;  
 RA Archer S., Hirano J., Diss J.K.J., Fraser S.P., Djangoz M.B.A.;  
 RT "Expression and localization of the fish retina of a homologue of the  
 RL Neuroreport 9:2049-2056 (1998).  
 DR EMBL: Y17128; CAA76641.1; -

DR	MEROPS: A22.001; -	
DR	InterPro: IPR001108; PresentIn.	
DR	Pfam: PF01080; PresentIn: 1.	
DR	PRINTS: PR01072; PresentIn.	
FT	NON_TER	1
FT	VARIANT	5
FT	VARIANT	56
FT	VARIANT	56
FT	VARIANT	56
FT	VARIANT	60
FT	VARIANT	63
FT	VARIANT	63
FT	VARIANT	83
FT	VARIANT	118
FT	VARIANT	200
Q	SEQUENCE	364 AA; 43276 MW; F97EEF2AB31FDAACRC64

Query Match	12.6%	Score 59	DB 13	Length 384
Best Local Similarity	100.0%	Pred. No.	2.5e-51	
Matches 59	Conservative 0	Mismatches 0	Indels 0	Gaps 0

```

Oy | 392 LVGKASATASGDWNTTTACFVAIIIGCLCTLLLLAIFKKALPALPISTPGLVFEYFATD 4500
    | |||||
Db | 309 LVGKASATASGDWNTTTACFVAIIIGCLCTLLLLAIFKKALPALPISTPGLVFEYFATD 3677

```

RESULT	7			
09M6T7		PRELIMINARY;	PRT;	456 AA.
ID	09M6T7			
AC	09M6T7;			
DT	01-NOV-1999 (TREMBLrel. 12, Created)			
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	Presentinlin-1.			
GN	PSEN.			
OS	Brachydanio rerio (Zebrafish) (Zebra danio).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;			
OC	Cyprinidae; Danio.			
OX	NCBI_TaxID:7955;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Leimer U., Haass C.;			
RT	"Expression, proteolytic processing, and amyloidogenic activity of			
RT	zebrafish (Danio rerio) presentinlin-1."			
RL	Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AJ132931; CAB40386.1; ..			
DR	MEROPS; A22.001; ..			
DR	ZFIN: ZDB-GENE-991119-4; psen.			
DR	InterPro: IPR001108; Presentinlin.			
DR	Pfam: PF01080; Presentinlin.1			
DR	PRINTS; PR01072; Presentinlin.			
QO	SEQUENCE 456 AA; 50981 MW; B90CA2280C874F8D CRC64;			

Query Match	9.08;	Score 42;	DB 13;	Length 456;
Best Local Similarity	100.08;	Pred. No. 6.4e-34;		
Matches 42;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Oy		409	ACFVAIIIGLCITLLLIATFKKALPALPISITFGLVFYFATD	450
Dbl		398	ACFVAIIIGLCITLLLIATFKKALPALPISITFGLVFYFATD	439

RESULT 8		
091991		
ID	PRELIMINARY:	PRT: 441 AA.
AC	091991,	
DT	01-OCT-2000 (TREMBLrel. 15, Created)	
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)	
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	
DE	Putative presenilin2 homolog.	
GN	PRE2.	
OS	Brachydanio rerio (zebrafish) (zebra danio).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;	

OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lardelli M., McCarly R., Tamme R., Kortschak D.:  
RT "The function of presenilin2 in zebrafish embryogenesis."  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: A0271195; CAB71930.1; -.  
DR MEROPS: A02\_002; -.  
DR InterPro: IPR001108; Presenilin.  
DR Pfam: PF01080; Presenilin.1.  
DR PRINTS: PR01072; PRESENILIN.  
SQ SEQUENCE 441 AA: 49169 MW: BC720F3398BE5F82A CRC64;

Query Match 8.4%; Score 39; DB 13; Length 441;  
Best Local Similarity 100.0%; Pred. No. 7.2e-31;  
Matches 39; Conservative 0; Mismatches 0; Indels 0;

Qy 409 ACEVAILIGLCITLLLLAIFKKALPALPISITFGVYFE 447  
|||||  
Db 383 ACEVAILIGLCITLLLLAIFKKALPALPISITFGVYFE 421

RESULT	9
Q90ZE4	
ID	Q90ZE4; PRELIMINARY; PRT: 441 AA.
AC	Q90ZE4;
DT	01-DEC-2001 (TREMBLrel. 19, Created)
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE	ps2.
GN	ps2.
OS	Brachydanio rerio (zebrafish) (zebra danio).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
OC	Cyprinidae; Danio.
OX	NCBI_TaxID=7955;
RN	[1]
RP	SEQUENCE FROM N.A.
RT	Jiang Y.-J., Davies A., Lewis J.;
RA	"The cloning and function of ps2 in zebrafish.";
RL	Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF176539; AAK61828.1; -
DR	InterPro; IPR001108; Pfsenlinin.
DR	Pfam; PF01080; Pfsenlinin; 1.
QO	SEQUENCE 441 AA; 49132 MW; B0080F2ABC6764BD CRC64;

Query Match	8.4%	Score 39;	DB 13;	Length 441;
Best Local Similarity	100.0%	Pred. No. 7.2e-31;		
Matches 39; Conservative	0;	Mismatches	0;	Indels 0;
		Gaps	0;	

Dy	QY
383	409
ACFVAITLGLCLTLLLLAIFKKALPALPISITTEGLVFFYE	ACFVAITLGLCLTLLLLAIFKKALPALPISITTEGLVFFYE
422	447

RESULT	10			
Q8WS58				
ID	Q8WS58	PRELIMINARY;	PRT;	504 AA.
AC	Q8WS58;			
DT	01-MAR-2002 (TREMblrel. 20, Created)			
DT	01-MAR-2002 (TREMblrel. 20, Last sequence update)			
DT	01-JUN-2002 (TREMblrel. 21, Last annotation update)			
DE	Presentilin.			
GN	PS.			
OS	Branchiostoma floridae (Florida lancelet) (Amphioxus).			
OC	Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;			
OC	Branchiostoma.			
OX	NCBI_TaxID=7739;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=21590369; PubMed=11733140;			
RA	Marl\u00edne-Mir A.; Canestro C.; Gonz\u00e1lez-Duarte R.; Albaladejo R.;			

```
RT "Characterization of the amphioxus presentin gene in a high gene-
RT density genomic region illustrates duplication during the vertebrate
RT lineage."
RL Gene 279:157-164(2001).
DR EMBL: AF369891; AAL40416.1; JOINED.
DR EMBL: AF369890; AAL40416.1; JOINED.
DR InterPro: IPR000847; HTH_LYSR.
DR InterPro: IPR001108; Presentin.
DR Pfam: PF01080; Presentin.
DR PRINTS: PR01072; Presentin.
DR PROSITE: PS00044; HTH_LYSR_FAMILY; UNKNOWN.1.
SQ SEQUENCE 504 AA; 53443 MW; 9C3794D0302859B4 CRC64;

Query Match
Best Local Similarity 100.0%; Score 35; DB 5; Length 504;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 409 ACPVALIGLCTLLLAIFKKALPALPISITFGL 443
Db 446 ACPVALIGLCTLLLAIFKKALPALPISITFGL 480

RESULT 11
08WS59 PRELIMINARY; PRT; 525 AA.
ID 08WS59
AC 08WS59;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Presentin.
GN PS.
OS Branchiostoma floridae (Florida lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma
OX NCBI_Taxid=7739;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21590369; Pubmed=11733140;
RA Martinez-Mir A., Canestro C., Gonzalez-Duarte R., Alpalat R.;
RT "Characterization of the amphioxus presentin gene in a high gene-
RT density genomic region illustrates duplication during the vertebrate
RT lineage."
RL Gene 279:157-164(2001).
DR EMBL: AF369891; AAL40414.1; JOINED.
DR EMBL: AF369890; AAL40414.1; JOINED.
DR InterPro: IPR000847; HTH_LYSR.
DR InterPro: IPR001108; Presentin.
DR Pfam: PF01080; Presentin.
DR PRINTS: PR01072; Presentin.
DR PROSITE: PS00044; HTH_LYSR_FAMILY; UNKNOWN.1.
SQ SEQUENCE 525 AA; 57598 MW; 2B14CF77A80F07DE CRC64;

Query Match
Best Local Similarity 100.0%; Score 35; DB 5; Length 525;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 409 ACPVALIGLCTLLLAIFKKALPALPISITFGL 443
Db 467 ACPVALIGLCTLLLAIFKKALPALPISITFGL 501

RESULT 12
09D616 PRELIMINARY; PRT; 332 AA.
ID 09D616
AC 09D616;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 0 day neonate head cDNA, RIKEN full-length enriched library,
DE clone:4833416A15, full insert sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
```

```
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=HEAD;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aikawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schirni L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seiya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL: AK014706; BAB29514.1;
DR InterPro: IPR001108; Presentin.
DR Pfam: PF01080; Presentin.
DR PRINTS: PR01072; PRESENTIN.
SQ SEQUENCE 332 AA; 37588 MW; 5728A699675A7DAD CRC64;

Query Match
Best Local Similarity 100.0%; Score 28; DB 11; Length 332;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 253 ISYVDLVAVLCPKGPLRLVETAEGRNE 280
Db 259 ISYVDLVAVLCPKGPLRLVETAEGRNE 286

RESULT 13
096P32 PRELIMINARY; PRT; 390 AA.
ID 096P32
AC 096P32;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Presentin 2 transcript variant 2 (Fragment).
GN PSEN2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu Y., Hu X., Zhou Y., Peng X., Yuan J., Qiang B.;
RL Submitted (Sep-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF416718; AAL16812.1;
DR InterPro: IPR001108; Presentin.
DR Pfam: PF01080; Presentin.
FT NON_TER 390
SQ SEQUENCE 390 AA; 43727 MW; A62C791CED9DCFB8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 28; DB 4; Length 390;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 253 ISYVDLVAVLCPKGPLRLVETAEGRNE 280
Db 259 ISYVDLVAVLCPKGPLRLVETAEGRNE 286

RESULT 14
```

091VS3 ID 091VS3 PRELIMINARY; PRT; 448 AA.  
 AC 091VS3;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Presentin 2.  
 GN PSEN2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NC NCBL\_TaxID=10099;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BREAST TUMOR;  
 RA Strausberg R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC010403; AAI10403.1; -;  
 DR MGD: MGI:109284; Psen2.  
 DR InterPro: IPR001108; Presentin.  
 DR Pfam: PF01080; Presentin; 1.  
 DR PRINTS: PRO1072; PRESENTIN.  
 SQ SEQUENCE 448 AA; 49964 MW; 2ED1DE7AFC3E0581 CRC64;

Query Match 6.0%; Score 28; DB 11; Length 448;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-19;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 253 ISYDLVAVLCPCGRLMLVETAEERNE 280  
 DB 259 ISYDLVAVLCPCGRLMLVETAEERNE 286

## RESULT 15

090X07 ID 090X07 PRELIMINARY; PRT; 451 AA.  
 AC 090X07;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Presentin 2.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 NC NCBL\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Korade Mirnics Z., Kerjanov S., Lovelock J., Corey S.J.;  
 RT "Cloning of chicken presentin.",  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY043493; AAK95409.1; -;  
 DR InterPro: IPR001108; Presentin.  
 DR Pfam: PF01080; Presentin; 1.  
 DR PRINTS: PRO1080; Presentin.  
 SQ SEQUENCE 451 AA; 50496 MW; 534E6364C627EBB0 CRC64;

Query Match 4.9%; Score 23; DB 13; Length 451;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-14;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 71 EELTLKGAHVIMLFPVPLCM 93  
 DB 83 EELTLKGAHVIMLFPVPLCM 105

## RESULT 16

09GU38 ID 09GU38 PRELIMINARY; PRT; 582 AA.  
 AC 09GU38;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Presentin.  
 GN PS.

OS Helix lucorum.  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;  
 CC Helicaceae; Helicidae; Helix.  
 NC NCBL\_TaxID=31229;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Rogaev E.I., Riazanskaya N.N., Dvorianchikov G., Grigorenko A.P.,  
 RA Tyrsin O.;  
 RT "Presentin gene isolated from mollusk Helix lucorum."  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF197881; AAC28518.1; -;  
 DR InterPro: IPR001108; Presentin.  
 DR Pfam: PF01080; Presentin; 1.  
 DR PRINTS: PRO1072; PRESENTIN.  
 SQ SEQUENCE 582 AA; 63782 MW; 07934E7C452199F0 CRC64;

Query Match 4.3%; Score 20; DB 5; Length 582;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-11;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 257 DLVAVLCPCGRLMLVETAAQ 276  
 DB 294 DLVAVLCPCGRLMLVETAAQ 313

## RESULT 17

09S1K7 ID 09S1K7 PRELIMINARY; PRT; 397 AA.  
 AC 09S1K7;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Putative presentin.  
 GN ATG29900  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 NC NCBL\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RX MEDLINE=20083487; PubMed=10617197;  
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldguy T.V.,  
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,  
 RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,  
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
 RA Coppenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,  
 RA Salzberg S.L., Fraser C.M., Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana."  
 RL Nature 402:761-768(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA lin X.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC007113; AAD23630.1; -;  
 DR InterPro: IPR001108; Presentin.  
 DR Pfam: PF01080; Presentin; 1.  
 DR PRINTS: PRO1072; PRESENTIN.  
 SQ SEQUENCE 397 AA; 44011 MW; DE6804D6CA186783 CRC64;

Query Match 3.2%; Score 15; DB 10; Length 397;  
 Best Local Similarity 100.0%; Pred. No. 2e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 380 KIGLGDFIFYSVLVG 394  
 DB 313 KIGLGDFIFYSVLVG 327

## RESULT 18

Q9SDG2 PRELIMINARY; PRT; 478 AA.  
AC Q9SDG2;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE EST A0058092(S1536) corresponds to a region of the predicted gene.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPPONBARE;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC clone:PO038F12.";  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP000836; BA88182.1;  
DR InterPro: IPR001108; Presentin.  
DR Pfam: PF01080; Presentin.1.  
DR PRINTS: PRO1072; PRESENTIN.  
SQ SEQUENCE 478 AA; 50100 MW; BC151E1021C96D3 CRC64;

Query Match 3.28; Score 15; DB 10; Length 478;  
Best Local Similarity 100.0%; Pred. No. 2.4e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 KLGIDFFYSVLVG 394  
DB 394 KLGIDFFYSVLVG 408  
|||||

## RESULT 19

Q97UC3 PRELIMINARY; PRT; 334 AA.  
AC Q97UC3;  
DT 01-OCT-2001 (TREMBlrel. 18, Created)  
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)  
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)  
DE Hypothetical protein SSO3092.  
CN SSO3092.  
OS Sulfolobus solfataricus.  
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
OC Sulfolobus.  
OX NCBI\_TaxID=2287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 35092 / DSM 1617 / P2;  
RX MEDLINE-21332296; PubMed-11427726;  
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,  
RA Aweyer M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,  
RA De Moors A., Eruso G., Fletcher C., Gordon P.M.K.  
RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,  
RA Thi-Ngoc H.P., Redder P., Schenk M.B., Thieriault C., Tolstrup N.,  
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,  
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;  
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 334 AA; 37973 MW; 97622ED88EBF26BB CRC64;

Query Match 1.9%; Score 9; DB 17; Length 334;  
Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 LLLFFFSFI 180  
DB 198 LLLFFFSFI 206  
|||||

## RESULT 20

Q9PV06 PRELIMINARY; PRT; 64 AA.  
AC Q9PV06;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE RAG1 protein (Fragment).  
GN RAG1.  
OS Plecopterosus altivelis.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Osmeridae; Plecopterosus.  
OX NCBI\_TaxID=61084;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RAG1AYU;  
RX MEDLINE-93398697; PubMed-10468597;  
RA Venkatesh B., Ning Y., Brenner S.;  
RT "Late changes in spliceosomal introns define clades in vertebrate evolution.";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:10267-10271(1999).  
DR EMBL; AF137198; A054565.1;  
FT NON\_TER 1 1  
SQ SEQUENCE 64 AA; 7139 MW; 4A931243C256DD35 CRC64;

Query Match 1.7%; Score 8; DB 13; Length 64;  
Best Local Similarity 100.0%; Pred. No. 5.7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 419 CLTLLLA 426  
DB 11 CLTLLLA 18  
|||||

## RESULT 21

Q9PV05 PRELIMINARY; PRT; 67 AA.  
AC Q9PV05;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE RAG1 protein (Fragment).  
GN RAG1.  
OS Harpadon sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Aulopiformes; Synodontidae; Harpadontinae; Harpadon.  
OX NCBI\_TaxID=94310;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RAG1BDUCK;  
RX MEDLINE-93398697; PubMed-10468597;  
RA Venkatesh B., Ning Y., Brenner S.;  
RT "Late changes in spliceosomal introns define clades in vertebrate evolution.";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:10267-10271(1999).  
DR EMBL; AF137199; A054566.1;  
FT NON\_TER 1 1  
SQ SEQUENCE 67 AA; 7452 MW; 3FF12959F4C8D08A CRC64;

Query Match 1.7%; Score 8; DB 13; Length 67;  
Best Local Similarity 100.0%; Pred. No. 5.9;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 419 CLTLLLA 426  
DB 14 CLTLLLA 21  
|||||

RESULT 22  
Q93EF9 PRELIMINARY; PRT; 83 AA.  
ID O93EF9  
AC O93EF9  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Putative sec-independent protein translocase (Fragment).  
GN TATC.  
OS Helicobacter hepaticus.  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group.  
OC Helicobacter.  
OX NCBI\_TaxID=32025;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=3B1;  
RA Ge Z., Feng Y., Fox J.G.;  
RT "Helicobacter hepaticus genome: construction of an ordered cosmid library and sequence analysis of the selected genomic regions.";  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF358706; AAL16696.1; -;  
DR InterPro: IPR002033; Translocase.  
DR Pfam: PF00902; UPF0032; 1.  
FT NON\_TER 1  
FT 83  
SQ SEQUENCE 83 AA; 9174 MW; D887FA1218BD7F3D CRC64;  
Query Match 1.7%; Score 8; DB 2; Length 83;  
Best Local Similarity 100.0%; Pred. No. 7.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 65 QDEEDEE 72  
Db 66 QDEEDEE 73  
RESULT 23  
Q8QTD1 PRELIMINARY; PRT; 104 AA.  
ID O8QTD1  
AC O8QTD1  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE MSSV320.  
OS White spot syndrome virus (WSSV).  
OC Viruses; unclassified viruses.  
OX NCBI\_TaxID=92652;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TAIWAN;  
RX MEDLINE=20517548; PubMed=11062040;  
RA Tsai M.F., Yu H.T., Tzeng H.F., Leu J.H., Chou C.M., Huang C.J.,  
RT Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;  
RT "Identification and characterization of a shrimp white spot syndrome virus (WSSV) gene that encodes a novel chimeric polypeptide of cellular-type thymidine kinase and thymidylate kinase.";  
RT Virology 277:100-110(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TAIWAN;  
RX MEDLINE=21444071; PubMed=11853398;  
RA Chen L.L., Leu J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,  
RT Lo C.F., Kou G.H.;  
RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white spot syndrome virus and characterization of the motif important for targeting VP35 to the nuclei of transfected insect cells.";  
RT Virology 293:44-53(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TAIWAN;  
RA Lo C.-F., Kou G.-H.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF440570; AAL89188.1; -;  
SQ SEQUENCE 104 AA; 12296 MW; 4F59C6883BC7941F CRC64;  
Query Match 1.7%; Score 8; DB 12; Length 104;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 170 SLILLEFF 177  
Db 28 SLILLEFF 35  
RESULT 24  
Q8VAM0 PRELIMINARY; PRT; 105 AA.  
ID Q8VAM0  
AC Q8VAM0  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Msv265.  
OS White spot syndrome virus (WSSV).  
OC Viruses; unclassified viruses.  
OX NCBI\_TaxID=92652;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21548311; PubMed=11689662;  
RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;  
RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF332093; AAL33268.1; -;  
SQ SEQUENCE 105 AA; 12409 MW; 7812D728683BC794 CRC64;  
Query Match 1.7%; Score 8; DB 12; Length 105;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 170 SLILLEFF 177  
Db 28 SLILLEFF 35  
RESULT 25  
Q63071 PRELIMINARY; PRT; 139 AA.  
ID Q63071  
AC Q63071  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last annotation update)  
DE Brahma protein (Fragment).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SPRAGUE-DAWLEY;  
RX MEDLINE=97154526; PubMed=9001244;  
RA Gestlund Fairants A.K., Blomquist P., Kwon H., Wrange O.;  
RT "Glucocorticoid receptor-glucocorticoid response element binding RT stimulates nucleosome disruption by the SWI/SNF complex.";  
RL Mol. Cell. Biol. 17:895-905(1997).  
DR EMBL: X99723; CAA68062.1; -;  
FT NON\_TER 1  
FT 139  
SQ SEQUENCE 139 AA; 16768 MW; B1A6F8886C011560 CRC64;  
Query Match 1.7%; Score 8; DB 11; Length 139;  
Best Local Similarity 100.0%; Pred. No. 11;



Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 EODEEDE 71  
 |||||||  
 Db 3 EODEEDE 10

RESULT 26  
 08VOL7 PRELIMINARY; PRT: 288 AA.

AC 08VOL7;  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Putative ABC transporter permease protein B.  
 GN BATN1953.0RF9.  
 OS Brucella abortus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID=235;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=344;  
 RA Bricker B.J.;  
 RT "n1953, a new element from Brucella abortus";  
 RL EMBL: AF54951: AAL5931.1: -  
 DR EMBL: AF54951: AAL5931.1: -  
 DR InterPro: IPR000515: BPD\_transp.  
 DR Pfam: PF00528: BPD\_transp. 1.  
 DR PROSITE: PS00402: BPD\_TRANS\_PNN\_MEMBER: UNKNOWN\_1.  
 SQ SEQUENCE 288 AA; 31696 MW; B5C20EA20BDFD8E CRC64;

Query Match 1.7%; Score 8; DB 2; Length 288;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SLLELFF 177  
 |||||||  
 Db 27 SLLELFF 34

RESULT 27  
 09SLP4 PRELIMINARY; PRT: 314 AA.

AC 09SLP4;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Bulb chitinase-1.  
 GN TBCL.  
 OS Tulipa bakeri.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae; Tulipa.  
 OX NCBI\_TaxID=110455;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yamagami T., Tsutsumi K., Ishiguro M.;  
 RT "CDNA cloning, sequence, and expression of tulip bulb chitinase-1";  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL  
 HYDROLASES).  
 CC EMBL: AB035668: BAA88408.1: -  
 DR InterPro: IPR000677: 25\_Globulin  
 DR InterPro: IPR001579: Chitinase\_18/2.  
 DR InterPro: IPR001233: Glyco\_hydro\_18.  
 DR Pfam: PF00704: Glyco\_hydro\_18: 1.  
 DR PRINTS: PR00551: 2SGLOBULIN.  
 DR PROSITE: PS01095: CHITINASE\_18: 1.  
 KW Glycosidase; Hydrolase  
 SQ SEQUENCE 314 AA; 35090 MW; AE9F5D31522A03E7 CRC64;

Query Match 1.7%; Score 8; DB 10; Length 314;  
 Best Local Similarity 100.0%; Pred. No. 23;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 SSLLEFF 176  
 |||||||  
 Db 7 SSLLEFF 14

RESULT 28  
 08ZFL5 PRELIMINARY; PRT: 317 AA.

AC 08ZFL5;  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Putative exported protein.  
 GN YPOL66.  
 OS Yersinia pestis.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Yersinia.  
 OX NCBI\_TaxID=632;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CO-92 / BIOVAR ORIENTALIS;  
 RX MEDLINE=21470413; PubMed=11586360;  
 RA Parkhill J., Wren B.W., Thomson N.R., Hitchen P.G., Mungall K.L.,  
 Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,  
 Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
 Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,  
 Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
 Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;  
 RA "Genome sequence of Yersinia pestis, the causative agent of plague";  
 RL Nature 413:523-527(2001).  
 DR EMBL: AJ414149; CAC90506.1: -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 317 AA; 34974 MW; A032A0DF974481D4 CRC64;

Query Match 1.7%; Score 8; DB 16; Length 317;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 LIISLIL 173  
 |||||||  
 Db 9 LIISLIL 16

RESULT 29  
 09RR94 PRELIMINARY; PRT: 347 AA.

AC 09RR94;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Hypothetical protein DR2600.  
 GN DR2600.  
 OS Deinococcus radiodurans.  
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;  
 OC Deinococcaceae; Deinococcus.  
 OX NCBI\_TaxID=1299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RI;  
 RX MEDLINE=20036896; PubMed=10567266;  
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
 Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
 Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
 Yamateyan J.J., Lam P., McDonald L., Uitterback T., Zalewski C.,  
 Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
 Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
 Fraser C.M.;  
 RA "Genome sequence of the radioresistant bacterium Deinococcus  
 radiodurans RI";  
 RT Science 286:1571-1577(1999).

DR EMBL: AE002089; AAF12144.1; -.  
DR TIGR: DR2600; -.  
DR InterPro: IPR002965; P-rich\_extensn.  
DR PRINTS: PR01217; PRICEXTENS.  
KW Hypothetical protein: Complete proteome.  
SQ SEQUENCE 347 AA; 34908 MW; 1C3CEBA9C1AE4E4C CRC64;

Query Match 1.7%; Score 8; DB 16; Length 347;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 352 RSTPESRA 359  
DB 274 RSTPESRA 281

## RESULT 30

ID 060121 PRELIMINARY; PRT; 354 AA.

AC 060121;  
DT 01-AUG-1998 (TREMBlrel. 07, Created)  
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
GN Hypothetical 39.3 kDa protein Cl6G5.07C in chromosome II.  
GN SPBC16G5.07C.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RA Lyne M., Wood V., Rajandream M.A., Barrell B.G., Beck A.,  
RA Reinhardt R.;  
RU Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.  
DR EMBL: AL023554; CAI19027.1; -.  
DR InterPro: IPR001107; Band\_7.  
DR InterPro: IPR001972; Stomatlin.  
DR Pfam: PF01145; Band\_7; 1.  
DR SMART: SM00244; PHB; 1.  
DR PROSITE: PS01270; BAND\_7; FALSE-NEG.  
KW Hypothetical protein.  
SQ SEQUENCE 354 AA; 39274 MW; 66B09F6A12BDC30 CRC64;

Query Match 1.7%; Score 8; DB 3; Length 354;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 395 KASATASG 402  
DB 265 KASATASG 272

## RESULT 31

ID 077794 PRELIMINARY; PRT; 369 AA.

AC 077794;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Multidrug resistance protein MDRL (Fragment).  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9822958; PubMed=9568108;  
RA Steingold S.F., Sharp N.J., McEahan M.C., Hughes C.S., Dunn S.E.,  
RA Page R.L.;  
\*Characterization of canine MDRL mRNA: its abundance in drug resistant

RT cell lines and in vivo.";  
RL Anticancer Res. 18:393-400(1998).  
DR EMBL: AF092810; AAC62505.1; -.  
DR InterPro: IPR001140; ABCtransprtTM.  
DR Pfam: PF00664; ABC\_membrane; 1.  
FT NON\_TER 1  
FT 369 369  
SQ SEQUENCE 369 AA; 40894 MW; DA2EC6CBF3328DBE CRC64;

Query Match 1.7%; Score 8; DB 6; Length 369;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 420 LTLILLAI 427  
DB 300 LTLILLAI 307

## RESULT 32

ID 0966H2 PRELIMINARY; PRT; 379 AA.

AC 0966H2;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
GN Hypothetical protein R05G9.2.  
GN R05G9.2.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Pelodermata; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
RT investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA R Wilson.;  
RT "The sequence of C. elegans cosmid R05G9.";  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA Waterston R.;  
RT "Direct Submission.";  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC006678; AAK68392.1; -.  
DR InterPro: IPR001622; K+channel\_pore.  
SQ SEQUENCE 379 AA; 43265 MW; 551B265B97FFBD2 CRC64;

Query Match 1.7%; Score 8; DB 5; Length 379;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 413 ALLIGLCL 420  
DB 270 ALLIGLCL 277

## RESULT 33

ID 092PX3 PRELIMINARY; PRT; 392 AA.

AC 092PX3;  
DT 01-MAY-1999 (TREMBlrel. 10, Created)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
GN At2G18410 protein.  
AT2G18410.

OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RX MEDLINE=20083487; PubMed=10617197;  
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,  
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,  
 RA Cronin L.A., Shen M., VanAken S.E., Unam L., Tallon L.J., Gill J.E.,  
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
 RA Copenhaver G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,  
 RA Salzberg S.L., Fraser C.M., Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 thaliana.";  
 RL Nature 402:761-768(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Lin X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC006439; AAD15496.1; -  
 SQ SEQUENCE 392 AA: 43648 MW: 57505E9A96D9DCA CRC64;

Query Match 1.7%; Score 8; DB 10; Length 392;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 364 LSSIIILAG 371  
 |||||||  
 Db 44 LSSIIILAG 51

RESULT 34  
 09V4U2 PRELIMINARY; PRT; 519 AA.  
 AC 09V4U2;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE CG8690 protein.  
 GN CG8690  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Fandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer J.R., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-T., Andrews-Pfannkoch C., Baldwin D.,  
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,  
 RA Fodor C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Goddek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegian C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry B., Morris J., Mostreffi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003837; AAF59083.1; -  
 DR HSSP: P21332; 1UOK  
 DR FlyBase; FBgn0033297; CG8690.  
 DR InterPro; IPR000461; Alpha\_amyase.  
 DR Pfam; PF00128; alpha-amyase; 1.  
 SQ SEQUENCE 519 AA: 59513 MW: 01B36FB58CF9A57 CRC64;

Query Match 1.7%; Score 8; DB 5; Length 519;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 420 LTLILLAI 427  
 |||||||  
 Db 7 LTLILLAI 14

RESULT 35  
 052557 PRELIMINARY; PRT; 522 AA.  
 AC 052557;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Rifamycin efflux protein.  
 GN RIFP  
 OS Amycolatopsis mediterranei (Nocardia mediterranei).  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Pseudonocardineae; Pseudonocardaceae; Amycolatopsis.  
 NCBI\_TaxID=33910;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S699;  
 RX MEDLINE=98174059; PubMed=9512878;  
 RA August P.R., Tang L., Yoon Y.J., Ning S., Mueller R., Yu T.W.,  
 RA Taylor M., Hoffmann D., Kim C.G., Zhang X., Hutchinson C.R.,  
 RA Floss H.G.;  
 RT "Biosynthesis of the ansamycin antibiotic rifamycin: deductions from  
 the molecular analysis of the rif biosynthetic gene cluster of  
 Amycolatopsis mediterranei S699.";  
 RL Chem. Biol. 5:69-79(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S699;  
 RX MEDLINE=98165773; PubMed=9497318;  
 RA Kim C.G., Yu T.W., Fryhle C.B., Handa S., Floss H.G.;  
 RT "3-Amino-5-hydroxybenzoic acid synthase, the terminal enzyme in the  
 formation of the precursor of mc7N units in rifamycin and related  
 antibiotics.";  
 RL J. Biol. Chem. 273:6030-6040(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S699;  
 RX MEDLINE=21201076; PubMed=11278540;  
 RA Yu T.-W., Mueller R., Mueller M., Zhang X., Draeger G., Kim C.-G.,

RA Leistner E., Floss H.G.;  
 RT Mutational analysis and reconstituted expression of the biosynthetic  
 genes involved in the formation of 3-amino-5-hydroxybenzoic acid, the  
 starter unit of rifamycin biosynthesis in Amycolatopsis mediterranei  
 RT 5699."  
 RL J. Biol. Chem. 276:12546-12555(2001).  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S699;  
 RA August P.R., Tang L., Yoon Y.J., Ning S., Mueller R., Hutchinson C.R.,  
 Floss H.G.;  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 CC 1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 DR EMBL: AF040570; AAC0125.1; -  
 DR InterPro: IPR004638; Eflux\_Emb.  
 DR InterPro: IPR003662; sub.transpporter.  
 DR TRIGRAMS: TIGR00711; efflux\_Emb: 1.  
 DR PROSITE: PS00216; SUGAR\_TRANSPORT\_1; UNKNOWN\_1.  
 KW Transmembrane.  
 SQ SEQUENCE 522 AA; 53208 MW; F09E79A9579638EF CRC64;

Query Match 1.7%; Score 8; DB 2; Length 522;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 247 WLILAVIS 254  
 |||||  
 DB 16 WLILAVIS 23

RESULT 36  
 Q03982 PRELIMINARY; PRT; 538 AA.  
 ID 003982;  
 AC 003982;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE P-glycoprotein.  
 GN PGP-1.  
 OS Cricetus sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Cricetulus.  
 OX NCBI\_TaxID=10031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-LUNG;  
 RX MEDLINE-91154265; PubMed-1671863;  
 RA Devine S.E., Hussain A., Davide J.P., Melera P.W.;  
 RT Full length and alternatively spliced pgp-1 transcripts in multidrug  
 resistant Chinese hamster lung cells."  
 RL J. Biol. Chem. 266:4545-4555(1991).  
 CC 1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.  
 DR EMBL: M59252; AAA37003.1; -  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR001140; ABCTransporter.  
 DR InterPro: IPR003439; ABC\_transport.  
 DR Pfam: PF00664; ABC\_membrane: 1.  
 DR Pfam: PF00005; ABC\_tran: 1.  
 DR ProDom: PD000006; ABC\_transport: 1.  
 DR SMART: SM00382; AAA; 1.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
 KW ATP-binding; Transport.  
 SQ SEQUENCE 538 AA; 59433 MW; 09A82E1818AE09C3 CRC64;

Query Match 1.7%; Score 8; DB 11; Length 538;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 420 LTVLLAI 427  
 |||||  
 DB 116 LTVLLAI 123

RESULT 37  
 Q08569 PRELIMINARY; PRT; 563 AA.  
 ID 08569;  
 AC 08569;  
 DT 01-JUN-2002 (TReMBLrel. 21, Created)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Similar to SWI/SNF related, matrix associated, actin dependent  
 DE regulator of chromatin, subfamily a, member 4 (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC023186; AAH23186.1; -  
 FT NON\_TER 1  
 SQ SEQUENCE 563 AA; 65393 MW; D90F4D3A2802B70F CRC64;

Query Match 1.7%; Score 8; DB 11; Length 563;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 EODEEDE 71  
 |||||  
 DB 204 EODEEDE 211

RESULT 38  
 Q09718 PRELIMINARY; PRT; 581 AA.  
 ID 09718;  
 AC 09718;  
 DT 01-OCT-2001 (TReMBLrel. 18, Created)  
 DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE Predicted membrane protein.  
 GN CAC1775.  
 OS Clostridium acetobutylicum.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
 OC Clostridiales; Clostridiaceae; Clostridium.  
 OX NCBI\_TaxID=1488;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;  
 RX MEDLINE-21359325; PubMed-11466286;  
 RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,  
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hittl J., Wolf Y.I.,  
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,  
 RA Bennett G.N., Koonin E.V., Smith D.R.;  
 RT "Genome sequence and comparative analysis of the solvent-producing  
 bacterium Clostridium acetobutylicum."  
 RL J. Bacteriol. 183:4823-4838(2001).  
 DR EMBL: AE007686; AAK79740.1; -  
 DR InterPro: IPR000566; Lipocin\_cytfABP.  
 DR PROSITE: PS00213; LIPOCALIN; UNKNOWN\_1.  
 KW Complete proteome.  
 SQ SEQUENCE 581 AA; 64675 MW; 34A8C602E514E9EF CRC64;

Query Match 1.7%; Score 8; DB 16; Length 581;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 187 KTYNVAVD 194  
 |||||  
 DB 232 KTYNVAVD 239

RESULT 39  
 Q095052 PRELIMINARY; PRT; 628 AA.  
 ID 095052

```
AC 095052;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE BRG-1-human (Fragment).
GN SMARCA4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Brower A., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Ganes J.,
RA Danganan L., Ertter A., Christensen M., Georgescu A., Avila J., Liu S.,
RA Atlix C., Andreise T., Frankheim M., Amico-Keller G., Coefield J.,
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RA Olsen A.S., Carrano A.V.;
RT "Sequence analysis of an 800 kb region in 19p13.2 between PDE4A and
RT DIS554."
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006127; AAC97986.1; -.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF002271; helicase_C.1.
DR Pfam; PF00176; SNF2_N; 1.
DR SMART; SM00490; HELICC; 1.
KW ATP-binding; Helicase.
FT NON_TER 1
FT 628 628
SQ SEQUENCE 628 AA; 73264 MW; 367164900E01A45B CRC64;

Query Match 1.7%; Score 8; DB 4; Length 628;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 EDDDEDE 71
DB 442 EDDDEDE 449

RESULT 40
O8BN06 PRELIMINARY; PRT; 668 AA.
AC O8BN06;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Multidrug resistance p-glycoprotein (Fragment).
GN MDRI.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Uozumi K., Nakaichi M.;
RT "Canine MDR."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB066299; BAB83959.1; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001140; ABC_transporter.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran.1.
DR Prodom; PD000006; ABC_transporter.1.
DR SMART; SM00382; AAA_1.
DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.
FT NON_TER 1
FT 668 668
SQ SEQUENCE 668 AA; 73546 MW; F8CD80F7C934B31A CRC64;
```

```
Query Match 1.7%; Score 8; DB 6; Length 668;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 420 LTLILLAI 427
DB 608 LTLILLAI 615

RESULT 41
O8ROK1 PRELIMINARY; PRT; 749 AA.
AC O8ROK1;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical 87.3 kDa protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC026672; AAH26672.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT 749 749
SQ SEQUENCE 749 AA; 87280 MW; 1FBCB0561A55BA CRC64;

Query Match 1.7%; Score 8; DB 11; Length 749;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 EDDDEDE 71
DB 391 EDDDEDE 398

RESULT 42
O63928 PRELIMINARY; PRT; 1022 AA.
AC O63928;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Brg1 protein (Fragment).
GN SMARCA4 OR BRG1.
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94123856; PubMed=7904967;
RA Randazzo F.M., Khavari P., Crabtree G., Tankun J., Rossant J.;
RT "Brg1: a putative murine homologue of the Drosophila brama gene, a
RT homeotic gene regulator."
RL Dev. Biol. 161:229-242(1994).
DR EMBL; S68108; AAC60670.1; -.
DR MGD; MGI:88192; Smarca4.
DR InterPro; IPR001487; Bromodomain.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR000330; SNF2_N.
DR Pfam; PF00439; bromodomain.1.
DR Pfam; PF00271; helicase_C.1.
DR Pfam; PF00176; SNF2_N; 1.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO_1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
```

DR PROSITE; PS00633; BROMODOMAIN\_1; 1.  
 DR PROSITE; PS50014; BROMODOMAIN\_2; 2.  
 KW ATP-binding; Helicase.  
 FT NON\_TER 1  
 SQ SEQUENCE 1022 AA; 117518 MW; 24F9DB4CA110436 CRC64;

Query Match 1.7%; Score 8; DB 11; Length 1022;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 DEDEDEDE 71  
 |||||  
 DB 663 DEDEDEDE 670

## RESULT 43

ID 09TSU2 PRELIMINARY; PRT; 1163 AA.

AC 09TSU2; 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Multidrug resistance related (Fragment).  
 OS Felis silvestris catus (Cat).  
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
 OX NCBI\_TaxID=9685;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Okai Y., Nakamura N., Matushiro H., Kato H., Setoguchi A., Yazawa M.,  
 RA Okuda M., Watarai T., Hasegawa A., Tsujimoto H.;  
 RT "Molecular analysis of multidrug resistance in lymphoma cells in the  
 RT cat.";  
 RU Am. J. Vet. Res. 0:0-0(1999).  
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.  
 DR EMBL; AB029153; BA88701.1;  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR001140; ABCtransportTM.  
 DR InterPro: IPR003439; ABC\_transport.  
 DR Pfam: PF00664; ABC\_membrane; 2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR Prodom: PD000006; ABC\_transport; 2.  
 DR SMART; SM00382; AAA; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 2.  
 DR ATP-binding; Transport.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 1163 AA; 128511 MW; 1B5BA13776A93A26 CRC64;

Query Match 1.7%; Score 8; DB 6; Length 1163;  
 Best Local Similarity 100.0%; Pred. No. 73;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 420 LTLILLAI 427  
 |||||  
 DB 798 LTLILLAI 805

## RESULT 44

ID 060502 PRELIMINARY; PRT; 1169 AA.

AC 060502; 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE P-glycoprotein.  
 GN Pgp-1.  
 OS Cricetus sp.  
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Cricetus.  
 OX NCBI\_TaxID=10031;  
 RN 11

RP SEQUENCE FROM N.A.

RC TISSUE-LUNG;  
 RX MEDLINE=91154265; PubMed=1671863;  
 RA Devine S.E., Hussain A., Davide J.P., Melera P.W.;  
 RT "Full length and alternatively spliced ppg-1 transcripts in multidrug  
 RT resistant Chinese hamster lung cells.";  
 RL J. Biol. Chem. 266:4545-4553(1991).  
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.  
 DR EMBL; M59254; AAA37005.1;  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR001140; ABCtransportTM.  
 DR InterPro: IPR003439; ABC\_transport.  
 DR Pfam; PF00664; ABC\_membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR Prodom: PD000006; ABC\_transport; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 2.  
 KW ATP-binding; Transport.  
 SQ SEQUENCE 1169 AA; 128939 MW; 72E25B7CE29DC185 CRC64;

Query Match 1.7%; Score 8; DB 11; Length 1169;  
 Best Local Similarity 100.0%; Pred. No. 74;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 420 LTLILLAI 427  
 |||||  
 DB 747 LTLILLAI 754

## RESULT 45

ID 097140 PRELIMINARY; PRT; 1231 AA.

AC 097140; 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE Prespore-specific protein (Fragment).  
 GN PSpC.  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.  
 OX NCBI\_TaxID=44689;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Hsu Y.-S., Chang W.-T., Newell P.C., Gross J.D.;  
 RT "A repressor controlling cell-type specific expression in  
 RT Dictyostelium.";  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF104350; AAD16881.1;  
 FT NON\_TER 1231  
 FT NON\_TER 1231  
 SQ SEQUENCE 1231 AA; 138196 MW; 074956F6094F5A58 CRC64;

Query Match 1.7%; Score 8; DB 5; Length 1231;  
 Best Local Similarity 100.0%; Pred. No. 77;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 DEDEDEEL 73  
 |||||  
 DB 1043 DEDEDEEL 1050

## RESULT 46

ID 09JK64 PRELIMINARY; PRT; 1272 AA.

AC 09JK64; 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Multidrug resistance protein 1a.  
 GN PGY1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-WISTAR: TISSUE-LIVER;  
 RA Hooiveld G.J.E.J., Wlms J.W.J., Hagenbuch B., Jansen P.L.M.,  
 RA Meijer D.K.F., Muller M.;  
 RT "Cloning and functional characterization of the rat multidrug  
 RT resistance protein Mdr1a";  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.  
 DR EMBL: AF257746; AF69007.1; -;  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR001140; ABCtransportTM.  
 DR InterPro: IPR003439; ABC\_transportr.  
 DR Pfam: PF00664; ABC\_membrane; 2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR ProDom: PD000006; ABC\_transportr; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
 DR ATP-binding; Transport.  
 KW SEQUENCE 1272 AA; 140326 MW; B5C1DF89DFC986C5 CRC64;  
 SO  
 Query Match 1.7%; Score 8; DB 11; Length 1272;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 420 LLLLLLAI 427  
 Db 849 LLLLLLAI 856  
 RESULT 47  
 0924L6  
 ID 0924L6 PRELIMINARY; PRT; 1272 AA.  
 AC 0924L6;  
 DT 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE P-glycoprotein.  
 GN MDR1A.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY;  
 RA Yamazaki M., Leake B.F., Kim R.B.;  
 RT "Molecular Cloning of Rat Mdr1a cDNA";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF286167; AAK83023.1; -;  
 DR InterPro: IPR001140; ABCtransportTM.  
 DR InterPro: IPR003439; ABC\_transportr.  
 DR Pfam: PF00664; ABC\_membrane; 2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR ProDom: PD000006; ABC\_transportr; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; UNKNOWN 2.  
 SO SEQUENCE 1272 AA; 140343 MW; D9A17457362DA0ED CRC64;  
 Query Match 1.7%; Score 8; DB 11; Length 1272;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 420 LLLLLLAI 427  
 Db 849 LLLLLLAI 856  
 RESULT 48  
 046605  
 ID 046605 PRELIMINARY; PRT; 1280 AA.  
 AC 046605;  
 DT 01-JUN-1998 (TREMblrel. 06, Created)  
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)

DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Multidrug resistance p-glycoprotein.  
 GN MDR1.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-COLON;  
 RA Puel O., Lepage J.F., Alvinerie M., Gallier P., Pineau T.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.  
 DR EMBL: AF045016; AAC02113.1; -;  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR001140; ABCtransportTM.  
 DR InterPro: IPR003439; ABC\_transportr.  
 DR Pfam: PF00664; ABC\_membrane; 2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR ProDom: PD000006; ABC\_transportr; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
 DR ATP-binding; Transport.  
 KW SEQUENCE 1280 AA; 141524 MW; 762DD5AF4C73306 CRC64;  
 SO  
 Query Match 1.7%; Score 8; DB 6; Length 1280;  
 Best Local Similarity 100.0%; Pred. No. 80;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 420 LLLLLLAI 427  
 Db 858 LLLLLLAI 865  
 RESULT 49  
 002793  
 ID 002793 PRELIMINARY; PRT; 1285 AA.  
 AC 002793;  
 DT 01-JUL-1997 (TREMblrel. 04, Created)  
 DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Multidrug resistance protein-1.  
 GN MDR1.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Longley M., Crawford A.M.;  
 RA "Ovine mdr1 gene";  
 RT submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.  
 DR EMBL: U78609; AAB58489.1; -;  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR001140; ABCtransportTM.  
 DR InterPro: IPR003439; ABC\_transportr.  
 DR Pfam: PF00664; ABC\_membrane; 2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR ProDom: PD000006; ABC\_transportr; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
 DR ATP-binding; Transport.  
 KW SEQUENCE 1285 AA; 142020 MW; 90153E617C44856F CRC64;  
 SO  
 Query Match 1.7%; Score 8; DB 6; Length 1285;  
 Best Local Similarity 100.0%; Pred. No. 80;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 420 LLLLLLAI 427  
 Db 862 LLLLLLAI 869

RESULT 50  
002569 PRELIMINARY; PRT: 1726 AA.  
AC 002569; 01-NOV-1996 (Tremblere, 01, Created)  
DT 01-NOV-1996 (Tremblere, 01, Last sequence update)  
DT 01-JUN-2002 (Tremblere, 21, Last annotation update)  
DE Merozoite surface antigen 1 precursor.  
GN PV200.  
OS Plasmodium vivax (Strain Belem).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=31273;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BELEM:  
RX MEDLINE-91219506; PubMed-2023952;  
RA del Portillo H.A., Longacre S., Khouri E., David P.H.;  
RT "Primary structure of the merozoite surface antigen 1 of Plasmodium  
RT vivax reveals sequences conserved between different Plasmodium  
RT species";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:4030-4034(1991).  
DR EMBL; M60807; AAA63427.1;  
DR InterPro; IPR000561; EGF-like.  
DR Pfam; PF00008; EGF; 1.  
KW Signal.  
FT SIGNAL. 1 17 POTENTIAL.  
FT CHAIN 18 1726 MEROZOITE SURFACE ANTIGEN 1.  
SQ SEQUENCE 1726 AA; 194434 MW; BA0DCD8333C6C727 CRC64;  
Query Match 1.7%; Score 8; DB 5; Length 1726;  
Best local Similarity 100.0%; Pred. No. 1e+02;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 173 LFFFSFI 180  
|||||||  
DB 4 LFFFSFI 11

Search completed: February 6, 2003, 14:20:37  
Job time : 89.3506 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 14:14:41 ; Search time 17.4398 seconds  
(without alignments)  
1110.644 Million cell updates/sec

Title: US-09-689-159A-2

Perfect score: 467  
Sequence: 1 MTELPAPLSYFONAQWSEDN.....ATDYLVQPMQLAFHQFYI 467

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size: 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	366	78.4	467	1 PSN1_HUMAN	P49768 homo sapien
2	107	22.9	467	1 PSN1_MICMU	P79802 microcebus
3	104	22.3	468	1 PSN1_RAT	P97887 rattus norv
4	97	20.8	467	1 PSN1_MOUSE	P49787 mus musculu
5	95	20.3	478	1 PSN1_BOVIN	O9XE97 bos taurus
6	59	12.6	433	1 PSN1_XENLA	O12976 xenopus lae
7	28	6.0	445	1 PSN2_XENLA	P79801 microcebus
8	28	6.0	448	1 PSN2_HUMAN	P49810 homo sapien
9	28	6.0	448	1 PSN2_MOUSE	O61144 mus musculu
10	28	6.0	448	1 PSN2_RAT	O88777 rattus norv
11	28	6.0	449	1 PSN2_BOVIN	O9XE96 bos taurus
12	27	5.8	449	1 PSN2_XENLA	O12977 xenopus lae
13	22	4.7	541	1 PSN1_DROME	O02194 drosophila
14	18	3.9	444	1 PSN1_CAEEL	P52166 caenorhabdi
15	15	3.2	453	1 PSN1_ARATH	O64668 arabidopsis
16	8	1.7	103	1 YGB0_ECOLI	O46894 escherichia
17	8	1.7	176	1 P22P1_RAT	P22282 rattus norv
18	8	1.7	263	1 NGN2_MOUSE	P70447 mus musculu
19	8	1.7	707	1 DCOE_LEIDO	P27116 leishmania
20	8	1.7	1276	1 MDR1_CRIGR	P21447 cricetus
21	8	1.7	1276	1 MDR3_MOUSE	P21447 mus musculu
22	8	1.7	1280	1 MDR1_HUMAN	P08183 homo sapien
23	8	1.5	58	1 TEM1_RANME	P79876 rana tempor
24	7	1.5	61	1 TEMB_RANME	P79874 rana tempor
25	7	1.5	61	1 TEMB_RANME	P79874 rana tempor
26	7	1.5	65	1 GAE5_RANME	P80399 rana rugosa
27	7	1.5	66	1 RLXN_RANCA	P39084 rana catesb
28	7	1.5	72	1 ATP1_BACCA	P41015 bacillus ca
29	7	1.5	72	1 ATP1_BACCP3	P00845 bacillus ps
30	7	1.5	72	1 ATP1_BACST	P42011 bacillus st
31	7	1.5	106	1 YWJ0_CAEEL	P34485 caenorhabdi
32	7	1.5	120	1 R22A_YEAST	P05749 saccharomyc
33	7	1.5	167	1 LSPA_CHLTR	O84413 chlamydia t

34	7	1.5	169	1 YFJ6_YEAST	P43604 saccharomyc
35	7	1.5	170	1 Y129_METHA	O57593 methanococ
36	7	1.5	173	1 OLE1_ARATH	P29525 arabidopsis
37	7	1.5	175	1 OLE2_BRANA	P29111 brassica na
38	7	1.5	183	1 OLE5_BRANA	P29109 brassica na
39	7	1.5	208	1 US10_HSVB	P30817 simian herp
40	7	1.5	213	1 UBC_ASFM2	P25869 african swi
41	7	1.5	243	1 RADG_SYNY3	P52601 synchocyst
42	7	1.5	271	1 ISPE_THEMA	O9X183 thermotoga
43	7	1.5	280	1 YIT1_BACSU	P39803 bacillus su
44	7	1.5	286	1 MCM1_YEAST	P11746 saccharomyc
45	7	1.5	291	1 Z1PA_VIBCH	O9K122 vibrio chol
46	7	1.5	296	1 SAPC_ECOLI	O47624 escherichia
47	7	1.5	300	1 NKX1_HISBI	O46383 bison bison
48	7	1.5	321	1 D7_AEDAE	P18183 aedes aegypt
49	7	1.5	323	1 RAD1_SCHPO	P22133 schizosacch
50	7	1.5	326	1 YE09_SYNY3	P73594 synchocyst
51	7	1.5	339	1 CRTB_RHCCA	P17056 rhodobacter
52	7	1.5	340	1 Y100_ARCFU	O28474 archaeoglob
53	7	1.5	347	1 NU2M_DIDMA	P41305 didelphis m
54	7	1.5	348	1 NU2M_MACRO	P92660 macropus ro
55	7	1.5	350	1 GANA_ASPAC	P48842 aspergillus
56	7	1.5	351	1 PSBD_GALSI	P28253 galderia s
57	7	1.5	351	1 PSBD_ODOSI	P49478 odontella s
58	7	1.5	351	1 PSBD_PORPU	P51357 porphyra pu
59	7	1.5	352	1 PSBD_CHLYU	P56319 chlorella v
60	7	1.5	352	1 PSBD_CYAPA	P48079 cyanophora
61	7	1.5	352	1 PSBD_NEPOL	O9L100 nephrolepim
62	7	1.5	352	1 PSBD_PEA	P06006 pisum sativ
63	7	1.5	352	1 PSBD_SPIOL	P06005 pinacia ol
64	7	1.5	352	1 PSBD_SYNP2	P20898 synchococ
65	7	1.5	352	1 PSBD_SYNP7	P11005 synchococ
66	7	1.5	352	1 PSBD_STNY3	P09182 synchocyst
67	7	1.5	353	1 PSBD_ARATH	P56761 arabidopsi
68	7	1.5	353	1 PSBD_HORVU	P18489 hordeum vul
69	7	1.5	353	1 PSBD_MAIZE	P48184 zeamays (m
70	7	1.5	353	1 PSBD_MARPO	P06404 marchantia
71	7	1.5	353	1 PSBD_ORKSA	O9MT11 cenothea h
72	7	1.5	353	1 PSBD_ORKSA	P12095 oryza sativ
73	7	1.5	353	1 PSBD_PINTH	P41644 pinus thunb
74	7	1.5	353	1 PSBD_POPE	O9XQ88 populus del
75	7	1.5	353	1 PSBD_SECC	P10803 secale cere
76	7	1.5	353	1 PSBD_TOBAC	P06403 nicotiana t
77	7	1.5	353	1 PSBD_WHEAT	O36814 triticum ae
78	7	1.5	353	1 YD53_ARCFU	O28918 archaeoglob
79	7	1.5	354	1 PSBD_MESVI	O2MUN2 mesotigma
80	7	1.5	361	1 SEPI_DROME	P42207 drosophila
81	7	1.5	361	1 SEPI_HUMAN	O15019 homo sapien
82	7	1.5	361	1 SEPI_MOUSE	P42208 mus musculu
83	7	1.5	370	1 RECF_BACSU	P05651 bacillus su
84	7	1.5	375	1 GUN1_ROBSP	P23044 robiliarda
85	7	1.5	377	1 FDN_EMENT	O03134 emeticella
86	7	1.5	379	1 VPP_BPMU	O08858 bacterioph
87	7	1.5	380	1 NTG2_YEAST	P08214 saccharomyc
88	7	1.5	381	1 CYB_ANTMI	O63524 antechinus
89	7	1.5	381	1 CYB_ANTSM	O33865 antechinus
90	7	1.5	381	1 CYB_DASYI	O34339 dasyrus vl
91	7	1.5	381	1 CYB_MURLO	O35020 murletia lon
92	7	1.5	381	1 CYB_PLANE	O35675 planigale t
93	7	1.5	396	1 SX11_CHICK	P48445 gallus galli
94	7	1.5	397	1 PACA_BPPI	P28813 bacterioph
95	7	1.5	404	1 YVBF_BACSU	P74748 bacillus su
96	7	1.5	405	1 SPI2_HUMAN	O75880 homo sapien
97	7	1.5	409	1 NU4M_ASCSU	P24880 ascaris suu
98	7	1.5	409	1 PROA_MYCLE	O9C675 mycobacteri
99	7	1.5	411	1 YAL2_SCHPO	O09665 schizosacch
100	7	1.5	415	1 PROA_BACST	P39821 bacillus su
101	7	1.5	415	1 PROA_MCTU	P71921 bacillus su
102	7	1.5	419	1 RNBP_MOUSE	P82313 mus musculu
103	7	1.5	421	1 NXP2_CAEEL	O9XV68 caenorhabdi
104	7	1.5	434	1 T147_HUMAN	O60664 homo sapien
105	7	1.5	440	1 Y284_AOUAE	O66658 aquilex aeo
106	7	1.5	440	1 YGR4_YEAST	P53107 saccharomyc

107	7	1.5	454	1	MYC2_MARMO	P20389	marmota mon	180	6	1.3	78	1	VNS7_CVPEFS	P05992	porcine tra
108	7	1.5	454	1	MYC2_SPEBE	Q64210	spermophilu	181	6	1.3	87	1	DEFA_RAT	Q62713	rattus norv
109	7	1.5	460	1	MYCN_MARMO	Q61976	marmota mon	182	6	1.3	87	1	DEFA_RAT	Q92111	rattus norv
110	7	1.5	462	1	MYCN_MOUSE	P03966	mus musculus	183	6	1.3	88	1	SABP_SARPE	P31529	sarcophaga
111	7	1.5	462	1	MYCN_RAT	Q63379	rattus norv	184	6	1.3	88	1	YMS0_ARCFU	O28053	archaeoglob
112	7	1.5	464	1	MYCN_HUMAN	P04198	homo sapien	185	6	1.3	89	1	EF1B_METTH	O27734	methanobact
113	7	1.5	477	1	EXU2_DROPS	Q24617	drosophila	186	6	1.3	93	1	DEFA_RAT	Q62714	rattus norv
114	7	1.5	482	1	MTA1_ANASP	P70802	anaabaena sp	187	6	1.3	93	1	RL23_HELPY	P56048	helicobacte
115	7	1.5	498	1	VGLX_PRIV1	P07562	pseudorabie	188	6	1.3	100	1	URE3_RHIME	P42887	rhizobium m
116	7	1.5	507	1	IRX3_MOUSE	P81067	mus musculus	189	6	1.3	101	1	THRP_HUMAN	P20962	homo sapien
117	7	1.5	509	1	MTB1_BACST	P70886	bacillus st	190	6	1.3	101	1	VE4_HP41	P27953	human papil
118	7	1.5	512	1	SYM_MTCPN	P37591	mycoplasma	191	6	1.3	101	1	YG09_YEAST	P53116	saccharomyc
119	7	1.5	532	1	YABM_BACSU	P73555	bacillus su	192	6	1.3	102	1	AP42_RAT	P04638	rattus norv
120	7	1.5	535	1	HTRI_HALN1	P33741	halobacteri	193	6	1.3	102	1	HSP3_MOUSE	Q62100	mus musculu
121	7	1.5	535	1	HTRI_HALSA	P33955	halobacteri	194	6	1.3	104	1	HSP3_RAT	Q64256	rattus norv
122	7	1.5	555	1	GPI_CHIRE	Q9fpg6	chlamydomon	195	6	1.3	106	1	YF81_ARCFU	O28691	archaeoglob
123	7	1.5	556	1	DPOL_MHW6	P11992	woodchuck h	196	6	1.3	108	1	CSGC_SALTY	P55327	salmoneila
124	7	1.5	566	1	CC45_HUMAN	O75419	homo sapien	197	6	1.3	109	1	Y819_METUA	Q58329	methanococc
125	7	1.5	591	1	YNA8_YEAST	P42846	saccharomyc	198	6	1.3	110	1	CSGC_ECOLI	P52107	escherichia
126	7	1.5	599	1	CENB_MOUSE	P27790	mus musculu	199	6	1.3	110	1	YB5A_METUA	P81316	methanococc
127	7	1.5	606	1	CENB_CRIGR	P48988	cricetus	200	6	1.3	111	1	PRO2_BOVIN	O09430	bos taurus
128	7	1.5	630	1	YCF2_OBNVI	P31569	oenothera v	201	6	1.3	114	1	HMG4_CHITE	P40622	chironomus
129	7	1.5	632	1	YK65_CABEL	P46555	caenorhabdi	202	6	1.3	114	1	NLR2_LYCES	P27056	lycopersico
130	7	1.5	665	1	PTH2_STNY3	P73179	synchocyst	203	6	1.3	114	1	NLR2_TOBAC	Q03461	nicotiana t
131	7	1.5	687	1	WHIT_DROME	P10090	drosophila	204	6	1.3	116	1	VAMS_HUMAN	O95183	homo sapien
132	7	1.5	694	1	NUCL_CHICK	P15771	gallus gall	205	6	1.3	119	1	TBCA_SCHPO	O97703	schizosacch
133	7	1.5	713	1	NUCL_MESAU	P08199	mesocricetu	206	6	1.3	120	1	NU3C_LUPLU	P52765	lupinus jut
134	7	1.5	721	1	YCF2_OENPI	P31568	oenothera p	207	6	1.3	126	1	PAND_XYTRA	O99977	xytella fcs
135	7	1.5	726	1	FATA_VIBAN	P1461	vibrio angu	208	6	1.3	126	1	VEBC_RAT	Q35357	rattus norv
136	7	1.5	742	1	NEB1_HUMAN	Q9u1j8	homo sapien	209	6	1.3	127	1	NRPF_ECO57	O83355	escherichia
137	7	1.5	748	1	PT1P_ECOLI	P37177	escherichia	210	6	1.3	127	1	NRPF_ECOLI	P32711	escherichia
138	7	1.5	748	1	PT1P_ECOLI	P37178	salmoneila	211	6	1.3	127	1	TRX3_YEAST	P23712	saccharomyc
139	7	1.5	766	1	SMZ7_BRARE	O9YHX4	brachydanio	212	6	1.3	129	1	RS8_ECOLI	P02361	escherichia
140	7	1.5	830	1	JTP2_MOUSE	O9ere9	mus musculu	213	6	1.3	131	1	YECN_ECOLI	P76289	escherichia
141	7	1.5	879	1	DPOL_MHV1	P03160	woodchuck h	214	6	1.3	132	1	TFW2_HUMAN	O01629	homo sapien
142	7	1.5	883	1	DPOL_MHV8	P06275	woodchuck h	215	6	1.3	134	1	LY6F_MOUSE	P35460	mus musculu
143	7	1.5	884	1	DPOL_MHV59	P12899	woodchuck h	216	6	1.3	135	1	Y549_RICPR	O92401	ricketsia
144	7	1.5	884	1	DPOL_MHV7	P12899	woodchuck h	217	6	1.3	138	1	YCX1_CHLYU	O20118	chlorella v
145	7	1.5	884	1	DPOL_MHV81	P17396	woodchuck h	218	6	1.3	139	1	IGF_MTXGL	P22618	myxine glut
146	7	1.5	906	1	FOX2_CANTR	P22414	candida tro	219	6	1.3	139	1	PRO2_HUMAN	P35080	homo sapien
147	7	1.5	932	1	SYL_ARCFU	O30250	archaeoglob	220	6	1.3	139	1	PRO2_MOUSE	O9j1v2	mus musculu
148	7	1.5	1085	1	YAFR_SCHPO	O09663	schizosacch	221	6	1.3	139	1	PRO2_RAT	O9epc6	rattus norv
149	7	1.5	1152	1	CD45_MOUSE	P06600	mus musculu	222	6	1.3	141	1	LYSA_BPP2	P51759	bacterioph
150	7	1.5	1160	1	WDR7_HUMAN	O9Y4e6	homo sapien	223	6	1.3	144	1	PAZA_HUMAN	P14555	homo sapien
151	7	1.5	1166	1	CY46_RAT	O03343	rattus norv	224	6	1.3	145	1	CY2_HROSH	P00095	rhodobacter
152	7	1.5	1234	1	PIP3_HUMAN	O01970	homo sapien	225	6	1.3	145	1	EXSC_PSPA6	P26995	pseudomonas
153	7	1.5	1234	1	PIP3_MOUSE	P51432	mus musculu	226	6	1.3	146	1	CYTC_HUMAN	P01034	homo sapien
154	7	1.5	1255	1	CD45_RAT	P04157	rattus norv	227	6	1.3	146	1	CYTC_MACMU	O19092	macaca mula
155	7	1.5	1279	1	MDR3_HUMAN	P21439	homo sapien	228	6	1.3	146	1	CYTC_SAI5C	O19093	saimiri sci
156	7	1.5	1302	1	MDR4_DROME	Q00449	drosophila	229	6	1.3	146	1	YU51_MOUSE	P58283	mus musculu
157	7	1.5	1304	1	CD45_HUMAN	P08575	homo sapien	230	6	1.3	148	1	YOEY_BACSU	P54464	baclillus su
158	7	1.5	1444	1	ADP1_MYCGE	P20796	mycoplasma	231	6	1.3	151	1	HSPH_BRAJA	O86110	bradyrhizob
159	7	1.5	2004	1	MOZ_HUMAN	O92794	homo sapien	232	6	1.3	151	1	YG27_METUA	O59021	methanococc
160	7	1.5	2175	1	HMCT_DROME	P10180	drosophila	233	6	1.3	152	1	Y38A_MYCPN	P75213	mycoplasma
161	7	1.5	2349	1	TPR_HUMAN	P12270	homo sapien	234	6	1.3	152	1	YLVS_BPPHY	P13004	lactococcus
162	7	1.5	2647	1	FLNA_HUMAN	P21333	homo sapien	235	6	1.3	154	1	RNM6_BOVIN	P08904	bos taurus
163	7	1.5	3530	1	MY15_HUMAN	O9ukn7	homo sapien	236	6	1.3	156	1	YF60_HAETN	P44253	haemophilus
164	7	1.5	4128	1	PRKD_MOUSE	P97313	mus musculu	237	6	1.3	156	1	YMS6_CAFEL	P34502	caenorhabdi
165	6	1.3	19	1	FIBB_TARTE	P14539	tapirus ter	238	6	1.3	157	1	EMPI_HUMAN	P54849	homo sapien
166	6	1.3	23	1	CERB_CERCA	P36191	ceratilis c	239	6	1.3	157	1	MNBY_MAI2E	P27347	zea mays (m
167	6	1.3	43	1	PSBN_PORPU	P51324	porphyra pu	240	6	1.3	157	1	YBEX_SALTY	O87574	salmoneila
168	6	1.3	56	1	RL32_UREPA	O9pgf8	ureaplasma	241	6	1.3	158	1	YU95_MYCTU	O95574	mycobacteri
169	6	1.3	57	1	RL32_ANASP	O8YV6F	anaabaena sp	242	6	1.3	160	1	UI95_HUMAN	O955x1	homo sapien
170	6	1.3	58	1	SCM2_MESMA	O9n1i5	mesobuthus	243	6	1.3	160	1	UI95_MOUSE	O9gc12	mus musculu
171	6	1.3	61	1	CCMD_BRAJA	P30959	bradyrhizob	244	6	1.3	161	1	HM6L_WHEAT	P40621	triticultum ae
172	6	1.3	65	1	OMP_LOCOMI	P80045	locusta mlg	245	6	1.3	166	1	RL21_ENTHI	P38653	entamoeba h
173	6	1.3	65	1	RPON_PYRAB	O9v194	pyrococcus	246	6	1.3	169	1	CCAC_CAVPO	O35505	cavia porce
174	6	1.3	65	1	RPNL_PYRHO	O59298	pyrococcus	247	6	1.3	170	1	MIR3_MOUSE	O9wtv3	mus musculu
175	6	1.3	74	1	BRTA_RANTE	P82268	rana tempor	248	6	1.3	170	1	NUIC_SPIOL	O9m3j7	spinaecla ol
176	6	1.3	74	1	BRTB_RANTE	P19672	locusta mlg	249	6	1.3	170	1	UBC6_HUMAN	O99462	homo sapien
177	6	1.3	77	1	AKH3_LOCOMI	O58132	methanococc	250	6	1.3	170	1	Y019_BORBO	O51051	borrella bu
178	6	1.3	77	1	FER7_METUA	P01076	lycopersico	251	6	1.3	173	1	LMIP_BOVIN	P20274	bos taurus
179	6	1.3	77	1	MCP1_LYCES			252	6	1.3	173	1	LMIP_HUMAN	P55344	homo sapien

253	6	1.3	173	1	LMIP_MOUSE	P56563	mus musculus	326	6	1.3	232	1	FL3L_MOUSE	P49772	mus musculus
254	6	1.3	173	1	LMIP_RAT	P54825	rattus norv	327	6	1.3	232	1	SMV_MOUSE	P13675	mus musculus
255	6	1.3	174	1	FTRV_SPIO	P38365	spiniacia ol	328	6	1.3	234	1	NO1L_RHIFR	P33212	rhizobium f
256	6	1.3	174	1	GSPW_PSEAE	P25061	pseudomonas	329	6	1.3	234	1	NO1L_RHISN	P55712	rhizobium s
257	6	1.3	174	1	YR19_HUMAN	P09002	homo sapien	330	6	1.3	237	1	LMPI_CAEEL	Q11117	caenorhabdl
258	6	1.3	176	1	22P2_RAT	P22283	rattus norv	331	6	1.3	238	1	Y939_HAEIN	P44080	haemophilus
259	6	1.3	177	1	GPIX_HUMAN	P14770	homo sapien	332	6	1.3	238	1	YHUO_YEAST	P32820	saccharomyc
260	6	1.3	178	1	Y793_METJA	O58203	methanococc	333	6	1.3	239	1	ANR2_ONCVO	P21842	onchocerca
261	6	1.3	179	1	SP17_MONDO	O62771	monodelphis	334	6	1.3	239	1	ATPI_CYACA	Q91M31	cyatidlum c
262	6	1.3	181	1	PTH_CAMJE	O9J117	campylobact	335	6	1.3	239	1	RIM9_YEAST	O04734	saccharomyc
263	6	1.3	182	1	SECT_PARDE	P72179	paracoccus	336	6	1.3	240	1	HDBG_HUMAN	P51858	homo sapien
264	6	1.3	183	1	HMLL_HUMAN	O9UJ13	homo sapien	337	6	1.3	240	1	SNF7_YEAST	P39929	saccharomyc
265	6	1.3	184	1	DSBE_VIBCH	O9K9E9	vibrio chol	338	6	1.3	241	1	YEEO_HAEIN	P43985	haemophilus
266	6	1.3	184	1	PHB1_SYNPW	O08087	synecococc	339	6	1.3	242	1	TPIS_MYCFL	P48779	mycoplasma
267	6	1.3	184	1	PHB1_SYNPW	O02180	synecococc	340	6	1.3	243	1	TPRIS_MYCGE	P47672	mycoplasma
268	6	1.3	186	1	Y545_CHLMU	O9PKC1	chlamydia m	341	6	1.3	244	1	TPIS_MYCPLA	P47670	mycoplasma
269	6	1.3	189	1	PA21_HUMAN	O9BZM1	homo sapien	342	6	1.3	245	1	TPIS_MYCPN	P78010	mycoplasma
270	6	1.3	190	1	CC42_COLGL	O94103	colletotric	343	6	1.3	245	1	CTRA_BOVIN	P00766	bos taurus
271	6	1.3	193	1	PMR2_MOUSE	P42925	mus musculu	344	6	1.3	246	1	RRM2_HUMAN	O9U143	homo sapien
272	6	1.3	193	1	PMR2_MOUSE	O07066	rattus norv	345	6	1.3	246	1	YIAT_ECO57	P58224	escherichia
273	6	1.3	195	1	GREP_FRATU	P48204	franciselila	346	6	1.3	246	1	YIAT_ECOLI	P58224	escherichia
274	6	1.3	195	1	PACA_CLAMA	P48144	clarias mac	347	6	1.3	247	1	LANP_RAT	P37681	escherichia
275	6	1.3	199	1	HMG4_HUMAN	O15347	homo sapien	348	6	1.3	247	1	YCIC_ECOLI	P49911	rattus norv
276	6	1.3	199	1	MLO3_SCHPO	O09330	schizosacch	349	6	1.3	248	1	MYCB_ECOLI	P75883	escherichia
277	6	1.3	200	1	ACPD_ECO57	O8X3S9	escherichia	350	6	1.3	249	1	MCT1_CANFA	P21842	canis famli
278	6	1.3	200	1	ACPD_ECOLI	P41407	escherichia	351	6	1.3	249	1	PHAI_HUMAN	P39687	homo sapien
279	6	1.3	200	1	ACPD_SALTU	O8XfB4	salmonella	352	6	1.3	249	1	YJUP_ECOLI	P39298	escherichia
280	6	1.3	200	1	HMG1_CHICK	P36194	gallus gall	353	6	1.3	251	1	UYP_YEAST	P18562	saccharomyc
281	6	1.3	201	1	HMG4_CHICK	P40618	gallus gall	354	6	1.3	252	1	CRB1_BOVIN	P12610	escherichia
282	6	1.3	201	1	MAP2_SCHPO	O09180	schizosacch	355	6	1.3	253	1	CYS2_ECOLI	O84173	chlamydia t
283	6	1.3	201	1	PSS_METJA	O58609	methanococc	356	6	1.3	253	1	TRPA_CHLTR	O48870	cyatidlum c
284	6	1.3	201	1	RUVA_CLOPE	O8X113	clostridium	357	6	1.3	255	1	COX2_CYACA	O97P88	streptococc
285	6	1.3	202	1	IPPD_BOVIN	P07516	bos taurus	358	6	1.3	255	1	DABP_STRPN	O8X6X7	escherichia
286	6	1.3	203	1	YH23_MYXVL	O9G6B4	myxoma viru	359	6	1.3	257	1	NUDC_ECO57	P32664	escherichia
287	6	1.3	203	1	YGIW_HAEIN	P44272	haemophilus	360	6	1.3	257	1	NUDC_ECOLI	O8Z338	salmonella
288	6	1.3	203	1	YH14_YEAST	P38897	saccharomyc	361	6	1.3	257	1	NUDC_SALTU	O91945	salmonella
289	6	1.3	204	1	ARR4_ARATH	O9SHU5	arabidopsis	362	6	1.3	257	1	NUDC_SALTU	O91945	salmonella
290	6	1.3	205	1	YNP2_CAEEL	P34555	caenorhabdl	363	6	1.3	258	1	ELI_HUMAN	O91945	salmonella
291	6	1.3	206	1	HMG2_CHICK	P26584	gallus gall	364	6	1.3	258	1	PMGE_HUMAN	O91945	salmonella
292	6	1.3	207	1	CL21_ARATH	P94014	arabidopsis	365	6	1.3	258	1	PMGE_MOUSE	P07338	homo sapien
293	6	1.3	208	1	YHNN_ECOLI	P37616	escherichia	366	6	1.3	258	1	PMGE_MOUSE	P07338	homo sapien
294	6	1.3	209	1	CYB_RHIFE	O21298	rhinolophus	367	6	1.3	261	1	COX3_LATCA	P07952	oryctolagus
295	6	1.3	209	1	HMG2_MOUSE	P30681	mus musculu	368	6	1.3	261	1	COX3_SCYCA	O03170	latimeria c
296	6	1.3	209	1	GPPI_HUMAN	O9NX15	homo sapien	369	6	1.3	261	1	COX3_SOUAC	O79407	scylliothun
297	6	1.3	209	1	GPPI_MOUSE	O9SEW8	mus musculu	370	6	1.3	262	1	COX3_METSE	O92248	squalus aca
298	6	1.3	210	1	SODK_PENCH	O75007	penicillium	371	6	1.3	262	1	ROB2_AGRRH	O47491	metridium s
299	6	1.3	210	1	YH90_SYNY3	P72793	synecocyst	372	6	1.3	264	1	FLIR_SALTU	P54702	salmonella
300	6	1.3	211	1	CLUD_MOUSE	O9Z054	mus musculu	373	6	1.3	264	1	GUNS_ERMCA	P16630	erythria car
301	6	1.3	211	1	HM1X_HUMAN	O9UQV6	homo sapien	374	6	1.3	265	1	DCMR_METSP	P45876	methylobact
302	6	1.3	213	1	NORD_CHLMU	O9PBH4	chlamydia m	375	6	1.3	265	1	MUR1_VIBCH	O9Kv17	vibrio chol
303	6	1.3	213	1	URE1_HELMU	P50044	helicobacte	376	6	1.3	266	1	EL1_BOVIN	O28153	bos taurus
304	6	1.3	214	1	HMG1_BOVIN	P10103	bos taurus	377	6	1.3	266	1	EL1_PIG	P00772	bos taurus
305	6	1.3	214	1	HMG1_HUMAN	P09429	homo sapien	378	6	1.3	266	1	EL1_PIG	P00772	bos taurus
306	6	1.3	214	1	HMG1_PIG	P12682	sus scrofa	379	6	1.3	267	1	GLNR_SPRCO	O05943	streptomyces
307	6	1.3	215	1	GTHE_ARATH	O96266	arabidopsis	380	6	1.3	268	1	EGF5_HUMAN	P12034	homo sapien
308	6	1.3	215	1	UBC_ASFB7	P27949	african swi	381	6	1.3	269	1	FLIR_BORBU	O44907	borrelia bu
309	6	1.3	215	1	YC66_RICCN	O9Z657	ricketsia	382	6	1.3	270	1	YFDD_CABEL	O11191	caenorhabdl
310	6	1.3	215	1	YDHR_ECOLI	P77564	escherichia	383	6	1.3	273	1	FPG_LACIC	P42371	lactococcus
311	6	1.3	216	1	YK36_AOUAE	O67827	aquilex aeo	384	6	1.3	273	1	TYRO_TRISI	P55026	p-triox s/n
312	6	1.3	218	1	Y363_RICPR	O9Z9G6	ricketsia	385	6	1.3	275	1	BPHB_PSEPS	P08604	p-clis-2,3-d
313	6	1.3	219	1	GPH_RHOCA	O33512	rhodobacter	386	6	1.3	277	1	CYSR_ECOLI	P16701	escherichia
314	6	1.3	220	1	TEGP_HSEVE4	P18348	equine hecp	387	6	1.3	277	1	CYSR_SALTU	P41032	salmonella
315	6	1.3	222	1	NMCA_SYNP7	P29283	synecococc	388	6	1.3	278	1	KC2B_YEAST	P43639	saccharomyc
316	6	1.3	222	1	Y815_RICPR	O9ZC62	ricketsia	389	6	1.3	280	1	ANFD_HELGE	O68955	hellobacter
317	6	1.3	223	1	NARI_BACSV	P42177	baclillus su	390	6	1.3	281	1	POT1_ECOLI	P31136	escherichia
318	6	1.3	223	1	NMCA_BANSP	O05061	anabaena sp	391	6	1.3	281	1	YH71_AERPE	O9YB24	aeropyrum p
319	6	1.3	224	1	DCU_LYCES	O42463	lycopersico	392	6	1.3	283	1	DAPA_METHU	O26892	methanobact
320	6	1.3	225	1	NMCA_SYNY3	P33779	synecocyst	393	6	1.3	285	1	YOFB_BACSM	O34685	baclillus su
321	6	1.3	227	1	PSAF_CHLRE	P12356	chlamydomon	394	6	1.3	286	1	HTPX_PASMU	P57846	pasteurella
322	6	1.3	228	1	YMC6_EMENT	P03883	emeritella	395	6	1.3	287	1	RS2_MYCTU	Q10796	mycobacteri
323	6	1.3	229	1	EFTU_GLOVI	P50064	gloeobacter	396	6	1.3	289	1	ATPg_BACFI	P22482	baclillus fl
324	6	1.3	229	1	JHBE_MANSE	P15500	manduca sex	397	6	1.3	290	1	SGCG_MOUSE	P82348	mus musculu
325	6	1.3	230	1	HXAV_HUMAN	P31268	homo sapien	398	6	1.3	290	1	YI42_AOUAE	O67656	aquilex aeo



545	6	1.3	386	1	CVCB_PEA	P13919	pisum sativ	618	6	1.3	428	1	ENO_CHLPN	092766	chlamydia p
546	6	1.3	386	1	NUCM_TRYBB	P21301	trypanosoma	619	6	1.3	428	1	HDA3_CHICK	P65520	gallus gall
547	6	1.3	386	1	RNAL_SCHPO	P41391	schizosacch	620	6	1.3	428	1	MTT8_THETH	P29749	thermus the
548	6	1.3	387	1	ASPP_AEPAE	Q03168	aedes aegypt	621	6	1.3	428	1	SSR3_MOUSE	P20935	mus musculus
549	6	1.3	387	1	QIN_AVIS3	P56260	avian sarco	622	6	1.3	428	1	SSR3_MOUSE	P20936	rattus norv
550	6	1.3	388	1	YMG7_MYCTU	Q50695	mycobacteri	623	6	1.3	429	1	THRC_BUCAI	P72869	buchnera ap
551	6	1.3	389	1	YND5_YEAST	P53962	saccharomyc	624	6	1.3	431	1	CST1_HUMAN	P00504	homo sapien
552	6	1.3	390	1	UGAT_MOUSE	Q91068	mus musculus	625	6	1.3	432	1	VISA_YEAST	P40564	homo sapien
553	6	1.3	392	1	TEBB_STMYI	P29548	stylonychia	626	6	1.3	434	1	PRSA_YEAST	P33257	saccharomyc
554	6	1.3	392	1	YU1J_ECOLI	P39381	eschcherichia	627	6	1.3	434	1	RP54_RHOSH	Q01194	rhodobacter
555	6	1.3	393	1	HEMX_ECOLI	P09127	eschcherichia	628	6	1.3	435	1	ARRB_CABEL	P14165	caenorhabdi
556	6	1.3	393	1	TRFR_MOUSE	P21761	mus musculus	629	6	1.3	436	1	IF4B_YEAST	P74878	saccharomyc
557	6	1.3	393	1	XYLH_ECOLI	P37389	eschcherichia	630	6	1.3	436	1	NOSD_PSBST	P19417	saccharomyc
558	6	1.3	394	1	URTG_DESRO	P49150	desmodus ro	631	6	1.3	438	1	ALG3_HUMAN	P09265	homo sapien
559	6	1.3	394	1	YD2C_SCHPO	Q10260	schizosacch	632	6	1.3	439	1	TEA3_MOUSE	P70210	mus musculus
560	6	1.3	395	1	SDC_DROME	P49419	drosophila	633	6	1.3	440	1	PUR8_HELPJ	Q92kx2	helicobacte
561	6	1.3	396	1	TRT_DROME	P19351	drosophila	634	6	1.3	440	1	PUR8_HELPJ	P56468	helicobacte
562	6	1.3	396	1	UGAT_HUMAN	P78381	homo sapien	635	6	1.3	442	1	MEAL_CHICK	P50680	gallus gall
563	6	1.3	397	1	DP3B_MYCSM	P52851	mycobacteri	636	6	1.3	443	1	ARGA_ECOLI	P08205	eschcherichia
564	6	1.3	397	1	EFTU_PSEAE	P09591	pseudomonas	637	6	1.3	444	1	FLIT_CANCN	O05528	caulobacter
565	6	1.3	398	1	YAIP_ECOLI	Q47536	eschcherichia	638	6	1.3	444	1	YCXD_BACSU	Q08752	bacillus su
566	6	1.3	398	1	YUGN_ECOLI	P39338	eschcherichia	639	6	1.3	445	1	GNT1_HUMAN	P26572	homo sapien
567	6	1.3	400	1	ENPL_MESAU	P08712	mesocricetu	640	6	1.3	445	1	YGCS_ECOLI	Q46909	eschcherichia
568	6	1.3	400	1	YD01_METJA	Q58697	methanococc	641	6	1.3	446	1	GNTU_ECOLI	P46856	eschcherichia
569	6	1.3	402	1	LAG2_CAEEL	P45442	caenorhabdi	642	6	1.3	446	1	GRMD_HUMAN	Q9bq67	homo sapien
570	6	1.3	402	1	RNBP_PIG	P17560	sus scrofa	643	6	1.3	447	1	CASR_HUMAN	P32239	homo sapien
571	6	1.3	404	1	GITS_HAEIN	P45240	haemophilus	644	6	1.3	447	1	GNT1_MOUSE	P27808	mus musculus
572	6	1.3	405	1	YUHB_ECOLI	P39352	eschcherichia	645	6	1.3	447	1	GNT1_RABIT	P27115	mus musculus
573	6	1.3	405	1	YLF2_YEAST	P38746	saccharomyc	646	6	1.3	447	1	GNT1_RABIT	P27115	mus musculus
574	6	1.3	406	1	MYC_BRARE	P52160	brachydanio	647	6	1.3	447	1	TBB3_ORYSA	Q09335	rattus norv
575	6	1.3	406	1	OTE_DROME	P20240	drosophila	648	6	1.3	448	1	TRME_NEITM	Q91wv2	oryza sativ
576	6	1.3	406	1	SNX6_HUMAN	Q9unb7	homo sapien	649	6	1.3	448	1	TRME_NEITM	Q91wv2	neisseria m
577	6	1.3	407	1	ISDF_RHTLO	Q98mx3	r_1spcd/1spf	650	6	1.3	449	1	LAT_MYCTU	P96855	mycobacteri
578	6	1.3	408	1	EFTU_CHACO	P50371	chakra commi	651	6	1.3	450	1	ECFE_SALTY	Q82964	salmonella
579	6	1.3	408	1	GPT_CRIGR	P24140	cricetulus	652	6	1.3	450	1	ECFE_SALTY	Q82964	salmonella
580	6	1.3	408	1	GPT_CRILLO	P23338	cricetulus	653	6	1.3	450	1	GSHR_ECOLI	P06715	eschcherichia
581	6	1.3	408	1	GPT_HUMAN	Q9h3b5	homo sapien	654	6	1.3	451	1	FXGB_CHICK	Q09064	gallus gall
582	6	1.3	409	1	R23B_HUMAN	P54727	homo sapien	655	6	1.3	452	1	CASR_RABIT	P46657	oryctolagus
583	6	1.3	409	1	Y055_CAEEL	Q09251	caenorhabdi	656	6	1.3	452	1	CASR_RAT	P30553	rattus norv
584	6	1.3	410	1	CH60_EHRCN	Q34194	erlichia c	657	6	1.3	452	1	PH4H_DROME	P17276	drosophila
585	6	1.3	410	1	EFTU_MESVI	Q4mup0	mesostigma	658	6	1.3	453	1	CASR_CANFA	P30552	canis fami
586	6	1.3	410	1	GPT_MOUSE	P42867	mus musculus	659	6	1.3	453	1	ST24_YEAST	P71514	saccharomyc
587	6	1.3	410	1	Y457_METJA	Q57899	methanococc	660	6	1.3	454	1	CASR_BOVIN	P79266	bos taurus
588	6	1.3	411	1	MP62_LYPTI	P91753	lytechinus	661	6	1.3	454	1	YGDB_ECOLI	P37320	oryctolagus
589	6	1.3	411	1	PALL_HUMAN	Q9um63	homo sapien	662	6	1.3	457	1	ARLY_ECO57	Q6x730	eschcherichia
590	6	1.3	411	1	RAPS_FORCA	P09106	torpedo cal	663	6	1.3	457	1	ARLY_ECOLI	P11447	eschcherichia
591	6	1.3	412	1	TRFR_RAT	Q01717	rattus norv	664	6	1.3	458	1	ARLY_SALTY	Q82311	salmonella
592	6	1.3	413	1	LCAT_CHICK	P53760	gallus gall	665	6	1.3	458	1	ARLY_SALTY	Q82311	salmonella
593	6	1.3	414	1	NSRI_YEAST	P27476	saccharomyc	666	6	1.3	458	1	ME31_DROME	P41298	balaenopter
594	6	1.3	414	1	UTII_HUMAN	Q9hwf9	homo sapien	667	6	1.3	459	1	NU4M_BALMU	P24975	balaenopter
595	6	1.3	415	1	VEGC_MOUSE	P97953	mus musculus	668	6	1.3	459	1	NU4M_BALMU	P24975	balaenopter
596	6	1.3	416	1	PICI_MOUSE	Q62083	mus musculus	669	6	1.3	460	1	ACMI_MACMU	P56489	maccaca mula
597	6	1.3	416	1	RAGE_BOVIN	Q28173	bos taurus-	670	6	1.3	460	1	NU4M_SCYCA	O79410	scyllorhinu
598	6	1.3	416	1	WZXE_ECOLI	P27834	eschcherichia	671	6	1.3	460	1	NU4M_SCYCA	O79410	scyllorhinu
599	6	1.3	418	1	SSR3_HUMAN	P32742	homo sapien	672	6	1.3	461	1	MTSL_SALIN	Q92z45	squalus aca
600	6	1.3	419	1	P47B_CANBO	Q00319	candida boi	673	6	1.3	462	1	SYG_TREBA	083678	treponema p
601	6	1.3	420	1	ODO2_BUCAI	P57389	buchnera ap	674	6	1.3	463	1	ARLY_SACDO	P41906	saccharomyc
602	6	1.3	420	1	YO20_BACHD	Q9kxat	baecillus ha	675	6	1.3	463	1	ARLY_YEAST	P44076	saccharomyc
603	6	1.3	421	1	CHDE_MUCRO	P50325	mucor rouxi	676	6	1.3	463	1	LTV1_YEAST	P40708	saccharomyc
604	6	1.3	421	1	MTTA_THBAO	P14385	thermus aqu	677	6	1.3	465	1	VP19_HVEB	P28935	equine herp
605	6	1.3	421	1	TES_HUMAN	Q9ug18	homo sapien	678	6	1.3	467	1	TRF6_MOUSE	P79431	mus musculus
606	6	1.3	422	1	PGL1_ARATH	P49062	arabidopsis	679	6	1.3	467	1	V51K_BMYVF	P09514	beet weester
607	6	1.3	422	1	YIUC_BACSU	Q31754	baecillus su	680	6	1.3	468	1	F121_BOVIN	P50448	bos taurus
608	6	1.3	423	1	EAT5_CAEEL	Q27295	caenorhabdi	681	6	1.3	468	1	KSGT_BRANA	O04160	drassica na
609	6	1.3	423	1	HRP_PLABE	Q08168	plasmodium	682	6	1.3	468	1	PTMB_BACHD	P36606	schizosacch
610	6	1.3	423	1	P47A_CANBO	P21245	candida boi	683	6	1.3	468	1	PTMB_BACHD	Q36578	baecillus ha
611	6	1.3	423	1	SAT_ENTHI	O76156	entamoeba h	684	6	1.3	469	1	YUW5_CAEEL	P32271	caenorhabdi
612	6	1.3	424	1	PUR2_AOUAE	O66949	aquifex aeo	685	6	1.3	471	1	RNT1_YEAST	O02555	saccharomyc
613	6	1.3	425	1	CABL_CAEEL	Q93249	caenorhabdi	686	6	1.3	472	1	CIST_DAUCA	O80433	daucus caro
614	6	1.3	426	1	BRN0_CORGL	O06754	corynebacte	687	6	1.3	474	1	Y068_MYGE	P47314	mycoplasma
615	6	1.3	426	1	TEAL_HUMAN	P28347	homo sapien	688	6	1.3	475	1	NCAP_SYNV	P10550	sonchus yel
616	6	1.3	426	1	TEAL_MOUSE	P30051	mus musculus	689	6	1.3	475	1	SIM1_YEAST	P40472	saccharomyc
617	6	1.3	427	1	MYCN_SERCA	P26014	serinus can	690	6	1.3	476	1	RPAE_HAEIN	O05074	haemophilus

691	6	1.3	477	1	INGR_MOUSE	P15261	mus musculus	764	6	1.3	533	1	MVIN_RHITR	O05467	rhizobium t
692	6	1.3	477	1	SYG_RHILO	Q984x8	rhizobium 1	765	6	1.3	534	1	PSD3_HUMAN	O43342	homo sapien
693	6	1.3	481	1	DDX6_XENIA	P54824	xenopus lae	766	6	1.3	535	1	PCD1_MAIZE	O43250	zea mays (m
694	6	1.3	481	1	PGKH_TOBAC	O42961	nicotiana t	767	6	1.3	535	1	UDJ1_RAT	O64550	rattus norv
695	6	1.3	481	1	TRME_RALSO	O8y3t5	ralettonia s	768	6	1.3	541	1	CATN_ASCSU	P90682	ascaris suu
696	6	1.3	482	1	NUSA_BOBBU	O51740	borrella bu	769	6	1.3	541	1	CATN_RHOSH	P93647	rhodobacter
697	6	1.3	482	1	T2EA_YEAST	P36100	saccharomyc	770	6	1.3	542	1	EAAL_BOVIN	P46411	bos taurus
698	6	1.3	483	1	BCA_STRYL	P33569	streptomyce	771	6	1.3	542	1	EAAL_HUMAN	P43003	homo sapien
699	6	1.3	483	1	DDX6_HUMAN	P26196	homo sapien	772	6	1.3	542	1	NCAP_P12HT	P21737	human parat
700	6	1.3	483	1	DDX6_MOUSE	P54823	mus musculus	773	6	1.3	542	1	TCPE_CAEEL	P47709	caenorhabdt
701	6	1.3	483	1	IDH_RICCN	O921r7	rickettsia	774	6	1.3	542	1	TUUL_HUMAN	O00294	homo sapien
702	6	1.3	483	1	IDH_RICCN	O92drt0	rickettsia	775	6	1.3	543	1	EAAL_MOUSE	P56564	mus musculus
703	6	1.3	484	1	MORE_CLOPE	O8xjy9	clostridium	776	6	1.3	543	1	EAAL_RAT	P24542	rattus norv
704	6	1.3	487	1	CATN_STRCO	Q92598	streptomyce	777	6	1.3	543	1	SGEL_YEAST	P33779	saccharomyc
705	6	1.3	487	1	CIT7_ECOLI	P77405	escherichia	778	6	1.3	544	1	CH60_AGRTS	P30735	agrobacteri
706	6	1.3	488	1	NU2M_OENBE	P93401	oenothera b	779	6	1.3	544	1	ESTP_DROME	P18167	drosophila
707	6	1.3	490	1	CHE1_HUMAN	O00409	homo sapien	780	6	1.3	544	1	PAK3_HUMAN	O75914	homo sapien
708	6	1.3	490	1	YNC2_YEAST	P53972	saccharomyc	781	6	1.3	544	1	PAK3_MOUSE	O61036	mus musculus
709	6	1.3	491	1	YIU0_YEAST	P40459	saccharomyc	782	6	1.3	544	1	PAK3_RAT	O62829	rattus norv
710	6	1.3	492	1	CPAC_MOUSE	P56593	mus musculus	783	6	1.3	545	1	CH60_PARDE	O92462	paracoccus
711	6	1.3	493	1	CPBL_MOUSE	O05421	mus musculus	784	6	1.3	545	1	CH60_RHOCA	P95678	rhodobacter
712	6	1.3	493	1	CPBL_RABIT	P08682	oryctolagus	785	6	1.3	545	1	CH60_RHIME	P35469	rhizobium m
713	6	1.3	493	1	HSF4_HUMAN	O9uiv5	homo sapien	786	6	1.3	545	1	CH60_RHIME	O922q4	rhizobium m
714	6	1.3	494	1	ACHE_RAT	P09660	rattus norv	787	6	1.3	546	1	CH60_RHILV	P34939	rhizobium l
715	6	1.3	495	1	CPBL_BOVIN	O18963	bos taurus	788	6	1.3	546	1	CH61_RHOSH	P20110	rhodobacter
716	6	1.3	495	1	CPBL_PIG	P79383	sus scrofa	789	6	1.3	547	1	CITR_KLEPN	P52687	klebsiella
717	6	1.3	496	1	ERG1_CANAL	Q92206	candida alb	790	6	1.3	548	1	CH60_ZYMO	P48280	zymomonas m
718	6	1.3	496	1	PAI2_HUMAN	O9up88	homo sapien	791	6	1.3	548	1	ERF_HUMAN	P50548	homo sapien
719	6	1.3	496	1	PR4H_HUMAN	O13523	homo sapien	792	6	1.3	548	1	LACT_PLURA	O01579	phlebia rad
720	6	1.3	496	1	PR4H_MOUSE	O61136	mus musculus	793	6	1.3	550	1	CH60_EHRCH	O42382	ehrllichia c
721	6	1.3	496	1	PSD3_ANGA	O61470	anopheles g	794	6	1.3	551	1	CBX4_MOUSE	O55187	mus musculus
722	6	1.3	497	1	C711_ARATH	O9sab6	arabidopsis	795	6	1.3	551	1	CH60_COWRU	P48213	cowdria rum
723	6	1.3	497	1	LMP2_EBV	P13285	epstein-bar	796	6	1.3	551	1	ERF_MOUSE	P70459	mus musculus
724	6	1.3	498	1	CCS_CAPAN	O42435	capicuin an	797	6	1.3	552	1	NCAP_TPMV	O9w4d0	tupaia para
725	6	1.3	500	1	CPGB_RAT	P08683	rattus norv	798	6	1.3	553	1	YQIK_ECOLI	P77306	escherichia
726	6	1.3	500	1	ICL1_HUMAN	P05155	homo sapien	799	6	1.3	555	1	PERF_HUMAN	P14722	homo sapien
727	6	1.3	500	1	LCYB_TOBAC	O43578	nicotiana t	800	6	1.3	556	1	CD19_HUMAN	P15911	homo sapien
728	6	1.3	500	1	PSD3_ANOST	O9u548	anopheles s	801	6	1.3	557	1	ATRA_CLOAB	O32327	clostridium
729	6	1.3	501	1	CSMD2_DROME	O9vmt6	drosophila	802	6	1.3	558	1	CBX4_HUMAN	O00257	homo sapien
730	6	1.3	501	1	NU2C_MARPO	P06257	maranthia	803	6	1.3	558	1	CX1B_PARDE	P98062	paracoccus
731	6	1.3	501	1	PYCA_METJA	O58626	methanococc	804	6	1.3	558	1	YJ83_MYCTU	O10873	mycobacteri
732	6	1.3	503	1	CCS_CITSI	O9se80	citrus sine	805	6	1.3	560	1	YJ60_HORVU	O00531	hordeum vul
733	6	1.3	503	1	LEB3_PEA	P05692	pisum sativ	806	6	1.3	561	1	CEAC_ECOLI	P00645	escherichia
734	6	1.3	504	1	SH2B_MOUSE	Q02152	mus musculus	807	6	1.3	562	1	CH60_TRYBB	O37683	trypanosoma
735	6	1.3	504	1	ICL1_MOUSE	P97290	mus musculus	808	6	1.3	562	1	CH60_TRYCR	O95046	trypanosoma
736	6	1.3	505	1	TUB_RAT	O88808	rattus norv	809	6	1.3	563	1	ENV_BAEVM	P10269	baabon endo
737	6	1.3	506	1	DHAI_YEAST	P39517	saccharomyc	810	6	1.3	565	1	COX1_RHOSH	P33317	rhodobacter
738	6	1.3	507	1	CATN_PICAN	P30263	pichia angu	811	6	1.3	565	1	NPH1_CANFA	O93119	canis famli
739	6	1.3	507	1	GIR6_HUMAN	O9u9q3	homo sapien	812	6	1.3	571	1	CVCA_PEA	P13915	plum sativ
740	6	1.3	508	1	EPOR_HUMAN	P19235	homo sapien	813	6	1.3	572	1	EA2_MOUSE	P43006	mus musculus
741	6	1.3	509	1	C4A3_DROME	O9vms7	drosophila	814	6	1.3	573	1	EA2_MOUSE	P31596	rattus norv
742	6	1.3	509	1	NUOM_ECOLI	P31978	escherichia	815	6	1.3	574	1	BEA1_YEAST	P47113	saccharomyc
743	6	1.3	512	1	ATPA_RICPR	O50288	rickettsia	816	6	1.3	574	1	EA2_HUMAN	P43004	homo sapien
744	6	1.3	512	1	EMRY_ECOLI	P52600	escherichia	817	6	1.3	575	1	YWA2_BACSU	P45861	baecillus su
745	6	1.3	515	1	YAHF_ECOLI	P77187	saccharomyc	818	6	1.3	577	1	YG48_YEAST	P53075	saccharomyc
746	6	1.3	516	1	SMR3_YEAST	Q04174	saccharomyc	819	6	1.3	583	1	CH12_CANAL	P40953	candida alb
747	6	1.3	517	1	Y652_HUMAN	O75143	homo sapien	820	6	1.3	583	1	SCP3_SCHPO	P41000	schizosacch
748	6	1.3	518	1	ANPD_AZOVI	P16266	azotobacter	821	6	1.3	585	1	BLAR_STRAU	P18357	staphylococ
749	6	1.3	520	1	CL1A_HORSE	O46515	equus cabal	822	6	1.3	585	1	MECR_STRAAM	P26597	staphylococ
750	6	1.3	520	1	YD99_LISMO	O91738	listeria mo	823	6	1.3	586	1	SYD_BUCAT	P57401	buchnera ap
751	6	1.3	521	1	TYRR_ERWHE	O921b7	erwinia her	824	6	1.3	586	1	YCAO_ECOLI	P75838	escherichia
752	6	1.3	523	1	ASSY_ARATH	O9sxx3	arabidopsis	825	6	1.3	587	1	RGP1_HUMAN	P46600	homo sapien
753	6	1.3	524	1	HUNB_TRICA	O01791	tribolium c	826	6	1.3	587	1	TALA_BPDV	P13894	budgerigar
754	6	1.3	524	1	NAB2_HUMAN	O15742	homo sapien	827	6	1.3	589	1	RGP1_MOUSE	P46061	mus musculus
755	6	1.3	526	1	ANPD_RHOCA	O07933	rhodobacter	828	6	1.3	593	1	DR1L_HUMAN	O99956	homo sapien
756	6	1.3	526	1	CL1A_MOUSE	O9q282	mus musculus	829	6	1.3	593	1	SOWT_YEAST	P36150	saccharomyc
757	6	1.3	528	1	YEPO_YEAST	P40039	saccharomyc	830	6	1.3	593	1	US22_HCMVA	P09722	human cytom
758	6	1.3	529	1	Y632_CHLTR	O84637	chlamydia t	831	6	1.3	598	1	ABB1_RAT	P46933	rattus norv
759	6	1.3	530	1	YABD_SCHPO	O09812	schizosacch	832	6	1.3	598	1	CYLL_HUMAN	P35653	homo sapien
760	6	1.3	530	1	PR09_YEAST	P19736	saccharomyc	833	6	1.3	599	1	CENB_HUMAN	P07199	homo sapien
761	6	1.3	530	1	PSO3_MOUSE	P14685	mus musculus	834	6	1.3	599	1	VG58_HSV1	O00157	icatalurid h
762	6	1.3	530	1	UD12_HUMAN	P36509	homo sapien	835	6	1.3	599	1	YAOB_SCHPO	O10110	schizosacch
763	6	1.3	533	1	FCY2_YEAST	P17064	saccharomyc	836	6	1.3	599	1	YV06_CAEEL	O93655	caenorhabdt

837	6	1.3	600	1	SF08_YEAST	P41833 saccharomyc	910	6	1.3	687	1	CSTA_HELPJ	O92447 helicobacte
839	6	1.3	602	1	VE1_CRPVK	P03112 cottonitall	911	6	1.3	687	1	CSTA_HELPJ	P56190 helicobacte
839	6	1.3	603	1	LCB2_SCHPO	O09925 schizosacch	912	6	1.3	691	1	HSF_DROME	P22813 dtrosophila
840	6	1.3	604	1	VE1_HPV38	O80909 human papil	913	6	1.3	692	1	ANR6_HUMAN	O49294 homo sapien
841	6	1.3	604	1	YFIC_BACSU	P54719 bacillus su	914	6	1.3	695	1	APP2_MOUSE	O06335 mus musculu
842	6	1.3	606	1	ABD4_HUMAN	O14678 homo sapien	915	6	1.3	697	1	NANB_STRPN	O54727 streptococc
843	6	1.3	606	1	ACEA_MYCLE	P46831 mycobacteri	916	6	1.3	699	1	SRCH_HUMAN	P23327 homo sapien
844	6	1.3	607	1	ALBU_BOVIN	P02769 bos taurus	917	6	1.3	700	1	MR11_CHICK	O91427 homo sapien
845	6	1.3	608	1	ALBU_SHEEP	P14639 ovis aries	918	6	1.3	702	1	YCBY_ECOLI	O91427 homo sapien
846	6	1.3	608	1	ALBU_SHEEP	P14639 ovis aries	919	6	1.3	706	1	NCCL_MOUSE	P75864 escherichia
847	6	1.3	614	1	AL11_ARAHY	P49064 felis silve	920	6	1.3	711	1	NCCL_MOUSE	P09405 mus musculu
848	6	1.3	614	1	VAAL_DROME	P48602 dtrosophila	921	6	1.3	712	1	CM4C_HUMAN	O53902 streptomyc
849	6	1.3	614	1	VAAL_DROME	O27331 dtrosophila	922	6	1.3	712	1	NCCL_RAT	O08493 homo sapien
850	6	1.3	614	1	VA51_HAEIN	O57180 haemophilus	923	6	1.3	712	1	SM1A_CARTEL	P13383 rattus norv
851	6	1.3	615	1	HIC2_HUMAN	O961b3 homo sapien	924	6	1.3	714	1	CADD_MOUSE	O17330 ceenorhadi
852	6	1.3	615	1	VAAL2_HUMAN	P38607 homo sapien	925	6	1.3	715	1	IGAA_YERE	O94715 mus musculu
853	6	1.3	615	1	VAAT_AEDAE	O16109 aedes aegyp	926	6	1.3	716	1	NCCL_MOUSE	P84722 yersinia pe
854	6	1.3	617	1	ASMA_ECOLI	P28249 escherichia	927	6	1.3	716	1	EMPL_RABIT	P05510 neurospora
855	6	1.3	617	1	VAAL_BOVIN	P31404 bos taurus	928	6	1.3	725	1	YV76_YEAST	O18750 oryctolagus
856	6	1.3	617	1	VAAL_HUMAN	P38606 homo sapien	929	6	1.3	726	1	TRE_BIADI	P71444 saccharomyc
857	6	1.3	617	1	VAAL_MOUSE	P50516 mus musculu	930	6	1.3	730	1	FER1_DROME	O02942 blaberus di
858	6	1.3	617	1	VAAL_MOUSE	P50516 mus musculu	931	6	1.3	730	1	FER1_DROME	O07407 dtrosophila
859	6	1.3	617	1	VAAL_MOUSE	P50516 mus musculu	932	6	1.3	732	1	AT2N_ECOLI	P37617 escherichia
860	6	1.3	619	1	RECO_HAEIN	P71359 haemophilus	933	6	1.3	732	1	POK_DROME	O01842 dtrosophila
861	6	1.3	620	1	GG95_HUMAN	O08379 homo sapien	934	6	1.3	743	1	PLAK_HUMAN	P14923 homo sapien
862	6	1.3	620	1	UL32_HSYEB	P28952 equine herp	935	6	1.3	743	1	YCR6_YEAST	P25353 saccharomyc
863	6	1.3	621	1	PLAK_MOUSE	O02257 mus musculu	936	6	1.3	746	1	APP_RAT	P36633 rattus norv
864	6	1.3	622	1	YXDM_BACSU	P42424 bacillus su	937	6	1.3	747	1	DL12_HUMAN	O91427 homo sapien
865	6	1.3	624	1	MTN4_MOUSE	O89029 mus musculu	938	6	1.3	757	1	IGR7_HUMAN	P06971 escherichia
866	6	1.3	624	1	NKX3_RAT	O9epq0 rattus norv	939	6	1.3	763	1	AOC3_HUMAN	O94853 homo sapien
867	6	1.3	625	1	RPOC_ANASP	P22704 anabaena sp	940	6	1.3	763	1	APP2_HUMAN	O16853 homo sapien
868	6	1.3	630	1	MI43_HUMAN	O9nr34 homo sapien	941	6	1.3	764	1	PGR8_HUMAN	O06481 homo sapien
869	6	1.3	631	1	YBFS_BACSU	P39816 bacillus su	942	6	1.3	765	1	APP2_RAT	P15943 rattus norv
870	6	1.3	632	1	RECO_PASBU	O9el21 pasteurella	943	6	1.3	765	1	S23A_HUMAN	O15436 homo sapien
871	6	1.3	633	1	SOHR_RHIME	P15715 rhizobium m	944	6	1.3	767	1	S23B_HUMAN	O15437 homo sapien
872	6	1.3	635	1	PTMA_KLEPN	O94bm7 klebsiella	945	6	1.3	767	1	S23B_MOUSE	O94662 mus musculu
873	6	1.3	635	1	RPSD_RICPR	P33451 rickettsia	946	6	1.3	768	1	PURL_SYNY3	P72644 synechocyst
874	6	1.3	636	1	GYRB_THEMA	P77993 thermotoga	947	6	1.3	769	1	UR06_VZVD	O09302 varicella-z
875	6	1.3	637	1	PTMA_STRPU	P15552 strongyloce	948	6	1.3	776	1	SM4F_RAT	O92143 rattus norv
876	6	1.3	637	1	PTMA_ECOLI	P00550 escherichia	949	6	1.3	777	1	SM4F_MOUSE	O92143 mus musculu
877	6	1.3	639	1	P2B1_CARTEL	O42773 cryptococcu	950	6	1.3	777	1	YASB_SCHPO	O10146 schizosacch
878	6	1.3	641	1	HM38_CAEBL	O419720 caenorhadi	951	6	1.3	778	1	EXP7_STRPN	P33597 streptococc
879	6	1.3	643	1	PD44_RAT	P38659 rattus norv	952	6	1.3	778	1	RG12_MOUSE	O61193 mus musculu
880	6	1.3	652	1	NU5M_PODAN	P20679 rodospore a	953	6	1.3	779	1	GFEL_YEAST	P37020 saccharomyc
881	6	1.3	653	1	CIK4_HUMAN	P22459 homo sapien	954	6	1.3	780	1	STRN_MOUSE	O55106 mus musculu
882	6	1.3	653	1	GP11_SCHPO	O14357 schizosacch	955	6	1.3	781	1	STRN_RAT	P70482 rattus norv
883	6	1.3	654	1	CIK4_MOUSE	O61423 mus musculu	956	6	1.3	781	1	CYNB_HUMAN	P35222 homo sapien
884	6	1.3	654	1	CIK4_MOUSE	O28527 mustela put	957	6	1.3	781	1	CYNB_MOUSE	O02248 mus musculu
885	6	1.3	655	1	CIK4_RAT	P15385 rattus norv	958	6	1.3	781	1	CYNB_MOUSE	O94u82 rattus norv
886	6	1.3	656	1	Y691_RICPR	O9zcm2 rickettsia	959	6	1.3	781	1	CYNB_RAT	P76823 xenopus lae
887	6	1.3	657	1	Y023_NPVOP	O10282 ortygia pseu	960	6	1.3	783	1	K6PF_XENLA	P76823 xenopus lae
888	6	1.3	660	1	CI01_SQDAC	O73925 squalus aca	961	6	1.3	784	1	K6PF_ASERG	P76823 xenopus lae
889	6	1.3	660	1	PH0B_ECOLI	P06972 escherichia	962	6	1.3	785	1	YFEB_THEMA	O94w59 thermotoga
890	6	1.3	661	1	CI80_HUMAN	O99467 homo sapien	963	6	1.3	788	1	YFEB_THEMA	O94w59 homo sapien
891	6	1.3	661	1	Y182_SYNY3	O55774 synechocyst	964	6	1.3	795	1	ENPL_CHICK	P08110 gallus galli
892	6	1.3	666	1	MUR2_ENTHR	P39045 enterococcu	965	6	1.3	795	1	TRKA_HUMAN	P04629 homo sapien
893	6	1.3	666	1	NU5M_CHOCR	P48920 chondrus cr	966	6	1.3	797	1	CDB9_HUMAN	O94541 homo sapien
894	6	1.3	666	1	NU5M_CHOCR	P48920 chondrus cr	967	6	1.3	797	1	Y648_HSVNA	O01033 herpesviru
895	6	1.3	667	1	RCG1_YEAST	P38339 saccharomyc	968	6	1.3	801	1	BRD2_HUMAN	P22440 homo sapien
896	6	1.3	668	1	UL52_HCMVA	P16793 human cytom	969	6	1.3	805	1	SECA_MYCCE	P47318 mycoplasma
897	6	1.3	674	1	CA1A_CHICK	P08125 gallus galli	970	6	1.3	806	1	SECA_MYCCE	P47318 mycoplasma
898	6	1.3	675	1	COP1_ARATH	P43254 arabidopsis	971	6	1.3	809	1	AK2H_ECOLI	P00562 escherichia
899	6	1.3	675	1	ORI_COTJA	P23499 coturnix co	972	6	1.3	809	1	CI0H_TORCA	P35522 torpedocal
900	6	1.3	679	1	GR75_CRIGR	O35501 cricetulus	973	6	1.3	812	1	FAED_ECOLI	P06970 escherichia
901	6	1.3	679	1	GR75_HUMAN	P38646 homo sapien	974	6	1.3	813	1	ARM_MUSDO	O04453 musca domes
902	6	1.3	679	1	GR75_MOUSE	P38647 mus musculu	975	6	1.3	817	1	RRPO_CRY	P14595 cymbidium r
903	6	1.3	679	1	GR75_MOUSE	P48721 rattus norv	976	6	1.3	817	1	RRPO_CRY	P14595 cymbidium r
904	6	1.3	680	1	GFAL_HUMAN	O06210 homo sapien	977	6	1.3	820	1	MDL2_YEAST	P15962 tomato bush
905	6	1.3	680	1	YH2X_SCHPO	O73443 schizosacch	978	6	1.3	823	1	CN8A_MOUSE	O88502 mus musculu
906	6	1.3	681	1	VE1_HPV10	P36720 human papil	979	6	1.3	824	1	YIP2_HUMAN	O13387 homo sapien
907	6	1.3	683	1	ERGL_THEMA	O9x1y4 thermotoga	980	6	1.3	826	1	RS65_HUMAN	O43374 homo sapien
908	6	1.3	683	1	SC31_RAT	O64319 rattus norv	981	6	1.3	827	1	ATC2_RHIME	P58342 rhizobium m
909	6	1.3	685	1	PH0B_SALTY	O87656 salmonella	982	6	1.3	830	1	BGLS_BUFETI	P16084 butyribdri

983	6	1.3	833	1	IF2_PASMD	P57873	pasteurella
984	6	1.3	834	1	GYRA_CHLPN	Q9Z843	chlamydia p
985	6	1.3	837	1	PHLD_MOUSE	070362	mus musculus
986	6	1.3	841	1	BAC2_HUMAN	Q9BYV9	homo sapien
987	6	1.3	843	1	ARM_DROME	P18824	drosophila
988	6	1.3	843	1	CPSB_CAEEL	O17403	caenorhabdi
989	6	1.3	843	1	DPOL_HPAVT	Q05486	hepatitis b
990	6	1.3	847	1	CD22_HUMAN	P20273	homo sapien
991	6	1.3	848	1	HEX_VIBVU	Q04786	vibrio vuln
992	6	1.3	852	1	SRCH_RABIT	P38266	saccharomyc
993	6	1.3	852	1	NCKX_DROME	P16230	oryctolagus
994	6	1.3	856	1	MOTS_BACSU	Q9U680	drosophila
995	6	1.3	858	1	MOTS_BACSU	P49849	drosophila
996	6	1.3	859	1	YOD3_CAEEL	P09263	caenorhabdi
997	6	1.3	860	1	YG2D_YEAST	P53246	saccharomyc
998	6	1.3	864	1	MS14_MOUSE	Q99M43	mus musculu
999	6	1.3	865	1	E78A_DROME	P45447	drosophila
1000	6	1.3	865	1	VGLB_HSYMD	P18538	marek's dis

## ALIGNMENTS

RESULT 1

PSNL\_HUMAN STANDARD; PRT; 467 AA.

AC P49768; Q14762; Q15719; Q15720;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Presenilin 1 (PS-1) (S182 protein).  
 GN PSEN1 OR PSNL1 OR AD3 OR PSI.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND VARIANTS AD (ISOFORMS I-467 AND I-463).  
 RC TISSUE=Brain;  
 RA MEDLINE=95319502; PubMed=7596406;  
 RA Sherrington R., Rogaeve E.I., Liang Y., Rogaeve E.A., Levesque G.,  
 RA Ikeda M., Chl H., Lin C., Li G., Holman K., Tsuda T., Mar L.,  
 RA Ponchi J.-F., Buntl A.C., Montesi M.P., Sobel S., Rainero I.,  
 RA Pinesse L., Nee L., Chumakov I., Pollen D., Brookes A.,  
 RA Sanson P., Polinsky R.J., Masco W., da Silva H.A.R., Haines J.L.,  
 RA Perlick-Vance M.A., Tanzi R.E., Roses A.D., Fraser P.E.,  
 RA Rommens J.M., St George-Hyslop P.H.;  
 RT "Cloning of a gene bearing missense mutations in early-onset familial  
 RT Alzheimer's disease."  
 RT Nature 375:754-760(1995) *5/2/95*  
 RL [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS I-463 AND I-374).  
 RC TISSUE=Blood, and Brain;  
 RA MEDLINE=96193901; PubMed=8641442;  
 RA Sahara N., Yahagi Y.-I., Takagi H., Kondo T., Okochi M., Usami M.,  
 RA Shirasawa T., Mori H.;  
 RT "Identification and characterization of presenilin I-467, I-463 and  
 RT I-374."  
 RT FEBS Lett. 381:7-11(1996).  
 RL [3]  
 RP SEQUENCE FROM N.A.  
 RA Rowen L., Madan A., Qin S., Abbasi N., Dors M., Ratcliffe A.,  
 RA Madan A., Dickhoff R., Shaffer T., James R., Lasky S., Hood L.;  
 RT "Complete sequence of the gene for presenilin 1."  
 RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 RL [4]  
 RP SEQUENCE OF 1-113 FROM N.A.  
 RA Tsujimura A., Hashimoto-Gotoh T.;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.  
 RX MEDLINE=96160372; PubMed=8574969;  
 RA Kovacs D.M., Fausett H.J., Page K.J., Kim T.-W., Moir R.D.,

RA Merriam D.E., Hollister R.D., Hallmark O.G., Mancini R.,  
 RA Felsenstein K.M., Hyman B.T., Tanzi R.E., Masco W.;  
 RT "Alzheimer-associated presenilins 1 and 2: neuronal expression in  
 RT brain and localization to intracellular membranes in mammalian  
 RT cells."  
 RL Nat. Med. 2:224-229(1996).  
 RN [6]  
 RP PROCESSING.  
 RX MEDLINE=97317150; PubMed=9173929;  
 RA Podlasky M.B., Citron M., Amarante P., Sherrington R., Xia W.,  
 RA Zhang J., Diehl T., Levesque G., Fraser P., Haass C., Koo E.H.,  
 RA Seubert P., St George-Hyslop P., Teplow D.B., Selkoe D.J.;  
 RT "Presenilin proteolysis undergoes heterogeneous endoproteolysis between  
 RT Thr291 and Ala299 and occur as stable N- and C-terminal fragments in  
 RT normal and Alzheimer brain tissue."  
 RL Neurobiol. Dis. 3:325-337(1997).  
 RN [7]  
 RP FUNCTION, AND MUTAGENESIS OF MET-292.  
 RX MEDLINE=20014554; PubMed=10545183;  
 RA Steiner H., Romig H., Pesold B., Philipp U., Baader M., Citron M.,  
 RA Loetscher H., Jacobsen H., Haass C.;  
 RT "Amyloidogenic function of the Alzheimer's disease-associated  
 RT presenilin 1 in the absence of endoproteolysis."  
 RL Biochemistry 38:14600-14605(1999).  
 RN [8]  
 RP FUNCTION.  
 RX MEDLINE=20062913; PubMed=10593990;  
 RA Ray W.J., Yao M., Mumm J., Schroeter E.H., Safitig P., Wolfe M.,  
 RA Selkoe D.J., Kopan R., Goate A.M.,  
 RT "Cell surface presenilin-1 particulates in the gamma-secretase-like  
 RT proteolysis of Notch."  
 RL J. Biol. Chem. 274:36801-36807(1999).  
 RN [9]  
 RP FUNCTION, AND MUTAGENESIS OF ASP-257 AND ASP-385.  
 RX MEDLINE=99221485; PubMed=10206644;  
 RA Wolfe M.S., Xia W., Ostaszewski B.L., Diehl T.S., Kimberly W.T.,  
 RA Selkoe D.J.;  
 RT "Two transmembrane aspartates in presenilin-1 required for presenilin  
 RT endoproteolysis and gamma-secretase activity."  
 RL Nature 398:513-517(1999).  
 RN [10]  
 RP FUNCTION, AND MUTAGENESIS OF ASP-257 AND ASP-385.  
 RX MEDLINE=20359495; PubMed=10899933;  
 RA Berzowska O., Jack C., McLean P., Aster J.C., Hicks C., Xia W.,  
 RA Wolfe M.S., Kimberly W.T., Weinmaster G., Selkoe D.J., Hyman B.T.;  
 RT "Aspartate mutations in presenilin and gamma-secretase inhibitors both  
 RT impair notch proteolysis and nuclear translocation with relative  
 RT preservation of notch signaling."  
 RL J. Neurochem. 75:583-593(2000).  
 RN [11]  
 RP FUNCTION, AND MUTAGENESIS OF LEU-286.  
 RX MEDLINE=20283925; PubMed=10811883;  
 RA Kulic L., Walter J., Multhaup G., Teplow D.B., Baumeister R.,  
 RA Romig H., Capell A., Steiner H., Haass C.;  
 RT "Separation of presenilin function in amyloid beta-peptide generation  
 RT and endoproteolysis of Notch."  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:5913-5918(2000).  
 RN [12]  
 RP FUNCTION.  
 RX PubMed=11226248;  
 RA Bakli L., Marambaud P., Efthimiopoulos S., Georgakopoulos A., Wen P.,  
 RA Cui W., Shioi J., Koo E., Ozawa M., Friedrich V.L., Robakis N.K.;  
 RT "Presenilin-1 binds cytoplasmic epithelial cadherin, inhibits  
 RT cadherin/p120 association, and regulates stability and function of  
 RT the cadherin/catenin adhesion complex."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:2381-2386(2001).  
 RN [13]  
 RP INTERACTION WITH DELTA-2 CATEININ.  
 RX MEDLINE=99155075; PubMed=10037471;  
 RA Levesque G., Yu G., Nishimura M., Zhang D.M., Levesque L., Yu H.,  
 RA Xu D., Liang Y., Rogaeve E., Ikeda M., Duthie M., Murgolo N., Wang L.,  
 RA Vanderveere P., Bayne M.L., Strader C.D., Rommens J.M., Fraser P.E.,  
 RA St George-Hyslop P.;



RT "Presenilins interact with armadillo proteins including  
 RT neural-specific plakophilin-related protein and beta-catenin."  
 RL J. Neurochem. 72:999-1008(1999).  
 RN [14]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE-97029239; PubMed-8875251.  
 RA Cruts M., Hendriks L., Van Broeckhoven C.;  
 RT "The presenilin genes: a new gene family involved in Alzheimer disease  
 pathology."  
 RL Hum. Mol. Genet. 5:1449-1455(1996).  
 RN [15]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE-98180715; PubMed-9521418;  
 RA Cruts M., van Broeckhoven C.;  
 RT "Presenilin mutations in Alzheimer's disease."  
 RL Hum. Mutat. 11:183-190(1998).  
 RN [16]  
 RP VARIANTS AD THR-143 AND ALA-384.  
 RX MEDLINE-96177673; PubMed-8634711;  
 RA Cruts M., Backhovens H., Wang S.-Y., van Gassen G., Theuns J.,  
 RA De Jonghe C., Wehnert A., de Voecht J., de Winter G., Cras P.,  
 RA Bryland M., Datsen N., Weissenbach J., den Dunnen J.T., Martin J.-J.,  
 RA Hendriks L., Van Broeckhoven C.;  
 RT "Molecular genetic analysis of familial early-onset Alzheimer's  
 RT disease linked to chromosome 14q24.3."  
 RL Hum. Mol. Genet. 4:2363-2372(1995).  
 RN [17]  
 RP VARIANTS AD L-82; H-115; T-139; R-163; T-231; L-264; V-392 AND Y-410.  
 RX MEDLINE-96177673; PubMed-8634712;  
 RA Camplon D., Flaman J.-M., Brice A., Hannequin D., Dubois B.,  
 RA Martin C., Moreau V., Charbonnier F., Didierjean O., Tardieu S.,  
 RA Peret C., Puel M., Pasquier F., Le Doze F., Bellis G., Calenda A.,  
 RA Helly R., Martinez M., Mallet J., Bellis M., Clerget-Darpoux F.,  
 RA Agid Y., Frebourg T.;  
 RT "Mutations of the presenilin 1 gene in families with early-onset  
 RT Alzheimer's disease."  
 RL Hum. Mol. Genet. 4:2373-2377(1995).  
 RN [18]  
 RP VARIANTS AD VAL-260; VAL-285 AND VAL-392.  
 RX MEDLINE-95379971; PubMed-7651536.  
 RA Rogaev E.I., Sherrington R., Rogaeva E.A., Levesque G., Ikeda M.,  
 RA Liang Y., Chh H., Lin C., Holman K., Tsuda T., Mar L., Sorbi S.,  
 RA Nacmias B., Placentini S., Amaducci L., Chumakov I., Cohen D.,  
 RA Lannfelt L., Fraser P.E., Romenes J.M., St George-Hyslop P.H.;  
 RT "Familial Alzheimer's disease in kindreds with missense mutations in  
 RT a gene on chromosome 1 related to the Alzheimer's disease type 3  
 RT gene."  
 RL Nature 376:775-778(1995).  
 RN [19]  
 RP VARIANTS AD V-139; V-146; Y-163; T-267; A-280 AND G-280.  
 RX MEDLINE-96024664; PubMed-7550356.  
 RA Clark R.F., Hutton M., Fuldner R.A., Froelich S., Karran E.,  
 RA Talbot C., Crook R., Lendon C.L., Prihar G., He C., Korenblatt K.,  
 RA Martinez A., Wieg M., Busfield F., Behrens M.I., Myers A., Norton J.,  
 RA Morris J., Mehta N., Pearson C., Lincoln S., Baker M., Duff K.,  
 RA Zehr C., Perez-Tur J., Houlden H., Ruiz A., Ossa J., Lopez F.,  
 RA Arcos M., Madrigal L., Collinge J., Humphreys C., Asworth T.,  
 RA Sarter S., Fox N.C., Harvey R., Kennedy A., Roques P.K., Cline R.T.,  
 RA Phillips C.A., Venter J.C., Forsel L., Axelman K., Lilius L.,  
 RA Johnston J., Cowburn R., Vitanen M., Winblad B., Kosik K.S.,  
 RA Haltia M., Poyhonen M., Dickson D., Mann D., Neary D., Snowden J.,  
 RA Lantos P., Lannfelt L., Rossor M.N., Roberts G.W., Adams M.D.,  
 RA Hardy J., Goate A.M.;  
 RT "The structure of the presenilin 1 (S182) gene and identification of  
 RT six novel mutations in early onset AD families."  
 RL Nat. Genet. 11:219-222(1995).  
 RN [20]  
 RP VARIANTS AD PHE-96; ARG-163 AND THR-213.  
 RX MEDLINE-96310408; PubMed-8733303.  
 RA Kaneko K., Sato S., Sasaki Y., Yoshiwaki A., Nishiwaki Y., Takeda H.,  
 RA Tanabe H., Nishimura T., Li K., St George-Hyslop P.H., Miki T.,  
 RA Oghara T.;  
 RT "Three different mutations of presenilin 1 gene in early-onset

RT Alzheimer's disease families."  
 RL Neurosci. Lett. 208:195-198(1996).  
 RN [21]  
 RP VARIANT AD ASP-135.  
 RX MEDLINE-97369208; PubMed-9225696;  
 RA Crook R., Ellis R., Shanks M., Thal L.J., Perez-Tur J., Baker M.,  
 Query Match 78.4%; Score 366; DB 1; Length 467;  
 Best Local Similarity 99.8%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;  
 Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MTELPAPLSYFQNAQMSQDNHLSNVRSQNDNREOEHNDRSLGHPPLNSGRPOGNSR 60  
 DB 1 MTELPAPLSYFQNAQMSQDNHLSNVRSQNDNREOEHNDRSLGHPPLNSGRPOGNSR 60  
 QY 61 QVEDEDEDEDELTKYGAHVIMLEVPVTCMVVVVATIKSVSFYTRKKDQLYTPPTE 120  
 DB 61 QVEDEDEDEDELTKYGAHVIMLEVPVTCMVVVVATIKSVSFYTRKKDQLYTPPTE 120  
 QY 121 DTEYVGOALSHSLNAAIMISVIVMTLLVLYVRYKYIHAMLLISSLLFFESFI 180  
 DB 121 DTEYVGOALSHSLNAAIMISVIVMTLLVLYVRYKYIHAMLLISSLLFFESFI 180  
 QY 181 YLGEVFKYNYAVDYITVALLIMNLGVGMISIMHKGPLRQQAAYLIMISALMALVFTKY 240  
 DB 181 YLGEVFKYNYAVDYITVALLIMNLGVGMISIMHKGPLRQQAAYLIMISALMALVFTKY 240  
 QY 241 LPETAWLILAVISYVDLVAVLCPKGPLMLVETAOERNETLFPALIVSSTWVWLVNMAE 300  
 DB 241 LPETAWLILAVISYVDLVAVLCPKGPLMLVETAOERNETLFPALIVSSTWVWLVNMAE 300  
 QY 301 GDPKQRRVSNKYNNAESTERESODTVAENDGGFSEMEQORSHGPHRSTPESRAA 360  
 DB 301 GDPKQRRVSNKYNNAESTERESODTVAENDGGFSEMEQORSHGPHRSTPESRAA 360  
 QY 361 VOELSSSILAGEDPEERGVKLGIDFIFYSVLYGKASATASGDMWTTIACFAVAILIGLCL 420  
 DB 361 VOELSSSILAGEDPEERGVKLGIDFIFYSVLYGKASATASGDMWTTIACFAVAILIGLCL 420  
 QY 421 TLLLIATFKKALPALPISITGVLVYFATDYLVOPFMDQLAFHQFYI 467  
 DB 421 TLLLIATFKKALPALPISITGVLVYFATDYLVOPFMDQLAFHQFYI 467  
 RESULT 2  
 PSNL\_MICMU STANDARD: PRT; 467 AA.  
 ID PSNL\_MICMU  
 AC P79802;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Presenilin 1 (PS-1).  
 GN PSEN1 OR PSNL1 OR PSL.  
 OS Microcebus murinus (Lesser mouse lemur).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Euthera; Primates; Strepsirhini; Cheirogaleidae;  
 CC Microcebus.  
 OX NCBI\_TaxID=30608;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Brain;  
 RX MEDLINE-97079199; PubMed-8920931;  
 RA Calenda A., Weste-Frances N., Czech C., Pradier L., Bons N.,  
 RA Bellis M.;  
 RT "Molecular cloning, sequencing, and brain expression of the  
 RT presenilin 1 gene in Microcebus murinus."  
 RL Blochm. Biophys. Res. Commun. 228:430-439(1996).  
 CC -1- FUNCTION: MAY PLAY A ROLE IN INTRACELLULAR SIGNALING AND GENE  
 CC EXPRESSION OR IN LINKING CHROMATIN TO THE NUCLEAR MEMBRANE.  
 CC -1- SUBUNIT: PREDOMINANTLY HETERODIMER OF A N-TERMINAL (NTF) AND A  
 CC C-TERMINAL (CTF) ENDOPROTEOLYTICAL FRAGMENT (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. GOLGI AND



```

|||||
DB 64 EGDEEDELFLTKGAKVHMLFVPTLCMVVVATIKSVSYTRKDDGLITPTPEDEE 123
OY 124 TVGQALSHILNAAIMISVIYVMTLLVLYKRYCKVYHAMLI 167
Db 124 TVGQALSHILNAAIMISVIYVMTLLVLYKRYCKVYHAMLI 167

RESULT 4
PSNL_MOUSE STANDARD: PRT: 467 AA.
AC PA9769; Q9JLP9;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Presenilin 1 (PS-1) (S182 protein).
GN PSN1 OR PSNL1 OR AD3H.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090.

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95319502; PubMed=7596406;
RA Sherrington R., Rogeev E.I., Liang Y., Rogeeva E.A., Levesque G.,
RA Ikeda M., Chl H., Lin C., Li G., Holman K., Tsuda T., Mar L.,
RA Foncin J.-F., Bruni A.C., Montesi M.P., Sorbi S., Rahner I.,
RA Pinesse L., Nee L., Chumakov I., Pollen D., Brookes A.,
RA Samsen P., Polinsky R.J., Masco W., da Silva H.A.R., Haines J.L.,
RA Perleack-Vance M.A., Tanzi R.E., Roses A.D., Fraser P.E.,
RA Rommens J.M., St George-Hyslop P.H.;
RT "Cloning of a gene bearing missense mutations in early-onset familial
RT Alzheimer's disease."
RT Nature 375:754-760(1995).
RL [2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RX MEDLINE=97442406; PubMed=9295283;
RA Mitsuda N., Roses A.D., Vitek M.P.;
RT "Transcriptional regulation of the mouse presenilin-1 gene."
RL J. Biol. Chem. 272:23489-23497(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=SAM P8; TISSUE=Hippocampus;
RA Kumar V.B., Vyas K.C., Choudhary V., Franko M., Flood J.F.,
RA Morley J.E.;
RT "Molecular cloning and tissue distribution of presenilin-1 in
RT senescence accelerated mice (SAM P8) mice."
RL Submitted (May-1999) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: MAY PLAY A ROLE IN INTRACELLULAR SIGNALING AND GENE
CC EXPRESSION OR IN LINKING CHROMATIN TO THE NUCLEAR MEMBRANE.
CC MAY FUNCTION IN THE CYTOPLASMIC PARTITIONING OF PROTEINS.
CC -1- SUBUNIT: PREDOMINANTLY HETERODIMER OF A N-TERMINAL (NTF) AND A
CC C-TERMINAL (CTF) ENDOPROTEOLYTIC FRAGMENT (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. GOLGI AND
CC ENDOPLASMIC RETICULUM (BY SIMILARITY).
CC -1- PTM: PHOSPHORYLATED ON SERINE RESIDUES (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: LA3177; AAC42094.1; -
CC EMBL: AF007560; AAB72049.1; -
CC EMBL: AF149111; AAF73153.1; -
CC MEROPS: A22.001; -
CC MGD: MGI:1202717; Psen1.

```

```

DR InterPro: IPR001108; Presenilin.
DR Pfam: PF01080; Presenilin.1.
DR PRINTS: PR01072; PRESENILIN.
KW Transmembrane; Phosphorylation; Endoplasmic reticulum; Golgi stack.
RN CHAIN 1
FT CHAIN 299 467 298 PRESENILIN 1 NTF SUBUNIT (BY SIMILARITY).
FT CHAIN 1 82 103 PRESENILIN 1 CTF SUBUNIT (BY SIMILARITY).
FT DOMAIN 83 103 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 104 132 LUMENAL (POTENTIAL).
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 154 160 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 161 181 POTENTIAL.
FT TRANSMEM 182 194 LUMENAL (POTENTIAL).
FT TRANSMEM 195 215 POTENTIAL.
FT TRANSMEM 216 220 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 221 241 POTENTIAL.
FT TRANSMEM 242 243 LUMENAL (POTENTIAL).
FT TRANSMEM 244 264 POTENTIAL.
FT TRANSMEM 265 407 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 408 428 POTENTIAL.
FT TRANSMEM 433 453 POTENTIAL.
FT TRANSMEM 433 453 POTENTIAL.
FT SITE 291 292
FT SITE 292 293
FT SITE 292 293
FT VARIANT 9 9 S -> T (IN STRAIN SAM P8).
FT VARIANT 40 40 D -> E (IN STRAIN SAM P8).
FT VARIANT 67 67 E -> CM (IN STRAIN SAM P8).
FT VARIANT 196 196 V -> L (IN STRAIN SAM P8).
FT VARIANT 321 322 ER -> RRD (IN STRAIN SAM P8).
SQ SEQUENCE 467 AA; 52639 MW; D07215B4BD2549 CRC64;

Query Match 20.8%; Score 97; DB 1; Length 467;
Best Local Similarity 100.0%; Pred. No. 6.8e-86;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 213 IHMKGPLRLOQAYLYMISALMALVFPKYLPENTAMVILAVISYDLYVAVLCRKGPLRMVY 272
Db 213 IHMKGPLRLOQAYLYMISALMALVFPKYLPENTAMVILAVISYDLYVAVLCRKGPLRMVY 272
OY 273 ETTAQRNETLFPALYYSTWVLVYVMAEGDPAQRV 309
Db 273 ETTAQRNETLFPALYYSTWVLVYVMAEGDPAQRV 309

RESULT 5
PSNL_BOVIN STANDARD: PRT: 478 AA.
ID PSNL_BOVIN
AC 09XJ97;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Presenilin 1 (PS-1).
GN PSN1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Sahara N., Shirasawa T., Mori H.;
RT "Molecular cloning of bovine presenilin 1 gene."
RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: MAY PLAY A ROLE IN INTRACELLULAR SIGNALING AND GENE
CC EXPRESSION OR IN LINKING CHROMATIN TO THE NUCLEAR MEMBRANE. MAY
CC FUNCTION IN THE CYTOPLASMIC PARTITIONING OF PROTEINS (BY
CC SIMILARITY).
CC -1- SUBUNIT: PREDOMINANTLY HETERODIMER OF A N-TERMINAL (NTF) AND A
CC C-TERMINAL (CTF) ENDOPROTEOLYTIC FRAGMENT (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. GOLGI AND
CC ENDOPLASMIC RETICULUM (BY SIMILARITY).

```

[illegible]

```

RT differential expression in oogenesis and embryogenesis."
RL Biochem. Biophys. Res. Commun. 231:392-396(1997).
CC -I- FUNCTION: MAY PLAY A ROLE IN NEGATIVE REGULATION OF APOPTOTIC
CC CASCADES DURING OOGENESIS AND EMBRYOGENESIS, AND IN
CC DEVELOPMENTALLY MATURED TISSUES SUCH AS BRAIN TISSUE.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -I- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN OVARIES AND TO A LESSER
CC EXTENT IN TESTIS, INTESTINE, KIDNEY, BRAIN, EYE AND LUNG. WEAK
CC EXPRESSION IN LIVER AND HEART. PRESENT IN TRACE AMOUNTS IN
CC SKELETAL MUSCLE.
CC -I- DEVELOPMENTAL STAGE: ABUNDANT IN EARLY STAGES OF OOGENESIS. THE
CC EXPRESSION IS RAPIDLY REDUCED BETWEEN MEIOTIC MATURATION AND
CC FERTILIZATION STAGES.
CC -I- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D84427; BAA19570.1; -.
DR MEROPS; A22.001; -.
DR InterPro; IPR001108; Presenilin.
DR Pfam; PF01080; Presenilin; 1.
DR PRINTS; PRO1072; PRESENILIN.
KW Transmembrane; Glycoprotein.
FT TRANSMEM 48 68 POTENTIAL.
FT TRANSMEM 99 119 POTENTIAL.
FT TRANSMEM 127 147 POTENTIAL.
FT TRANSMEM 161 181 POTENTIAL.
FT TRANSMEM 184 204 POTENTIAL.
FT TRANSMEM 210 230 POTENTIAL.
FT TRANSMEM 247 267 POTENTIAL.
FT TRANSMEM 374 394 POTENTIAL.
FT TRANSMEM 399 419 POTENTIAL.
FT CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 433 AA; 48301 MW; 71CCEB3F6B93C0AF CRC64;

Query Match 12.6%; Score 59; DB 1; Length 433;
Best Local Similarity 100.0%; Pred. No. 3,9e-49;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 409 ACPVALITGLCTLTLLIAIFKKAIPALPISITFGIVFYFATDYLVQPMQDLAFHOFYI 467
Db 375 ACPVALITGLCTLTLLIAIFKKAIPALPISITFGIVFYFATDYLVQPMQDLAFHOFYI 433

RESULT 7
PSN2_MICMU
ID PSN2_MICMU STANDARD: PRT: 445 AA.
AC P79801;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUN-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Presenilin 2 (PS-2) (Fragment).
OS PSN2 OR PSN2 OR PS2.
CN Microcebus murinus (Lesser mouse lemur).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirrhini; Cheirogaleidae;
OC Microcebus.
OX NCBI_TaxID=30608;
RX [1]
RA SEQUENCE FROM N.A.
RA TISSUE=Brain;
RA MEDLINE=99167105; PubMed=10069575;
RA Calenda A., Mestre-Frances N., Czech C., Pradier L., Peller A.,
RA Perret M., Bons N., Bellis M.;
RA "Cloning of the presenilin 2 cDNA and its distribution in brain of
RA the primate, Microcebus murinus: coexpression with betaAPP and Tau

```

RT proteins.";  
 RL Neurobiol. Dis. 5:323-333(1998).  
 CC -1- FUNCTION: MAY PLAY A ROLE IN INTRACELLULAR SIGNALING AND GENE  
 CC EXPRESSION OR IN LINKING CHROMATIN TO THE NUCLEAR MEMBRANE. MAY  
 CC FUNCTION IN THE CYTOPLASMIC PARTITIONING OF PROTEINS (BY  
 CC SIMILARITY).  
 CC -1- SUBUNIT: PREDOMINANTLY HETERODIMER OF A N-TERMINAL (NTF) AND A  
 CC C-TERMINAL (CTF) ENDOPROTEOLYTICAL FRAGMENT (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. GOLGI AND  
 CC ENDOPASMIC RETICULUM (BY SIMILARITY).  
 CC -1- PTM: PHOSPHORYLATED ON SERINE RESIDUES (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Y10140: CAA71228.1; -.  
 DR MEROPS: A22.002; -.  
 DR InterPro: IPR001108: Presenilin.  
 DR Pfam: PF01080: Presenilin.1.  
 DR PRINTS: PR01072: PRESENILIN.  
 KM Transmembrane: Phosphorylation: Endoplasmic reticulum; Golgi stack.  
 FT CHAIN 1 297 PRESENILIN 2 NTF SUBUNIT (BY SIMILARITY).  
 FT CHAIN 1 297 PRESENILIN 2 CTF SUBUNIT (BY SIMILARITY).  
 FT DOMAIN 1 87 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 88 108 POTENTIAL.  
 FT DOMAIN 109 138 LUMENAL (POTENTIAL).  
 FT TRANSSEM 139 159 POTENTIAL.  
 FT DOMAIN 160 166 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 167 187 POTENTIAL.  
 FT DOMAIN 188 200 LUMENAL (POTENTIAL).  
 FT TRANSSEM 201 221 POTENTIAL.  
 FT DOMAIN 222 223 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 224 244 POTENTIAL.  
 FT DOMAIN 245 249 LUMENAL (POTENTIAL).  
 FT TRANSSEM 250 271 POTENTIAL.  
 FT DOMAIN 272 388 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 389 409 POTENTIAL.  
 FT TRANSSEM 414 434 POTENTIAL.  
 FT NON\_TER 445 445  
 FT SEQUENCE 445 AA: 49475 MW: 0A01A764659E052 CRC64;  
 SQ  
 Query Match 6.0%; Score 28; DB 1; Length 445;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-19;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RA Pettingell W.H., Yu C.-E., Jondro P.D., Schmidt S.D., Wang K.,  
 RA Crowley A.C., Fu Y.-H., Guenette S.Y., Galas D., Nemens E.,  
 RA Majumdar E.M., Bird T.D., Schellenberg G.D., Tanzi R.E.,  
 RT "Candidate gene for the chromosome 1 familial Alzheimer's disease  
 RT locus.";  
 RT Science 269:973-977(1995).  
 RL  
 RN  
 RP SEQUENCE FROM N.A., AND VARIANTS FAD ILE-141 AND VAL-239.  
 RC TISSUE=Brain, and Colon;  
 RX MEDLINE=95379971; PubMed=7651536;  
 RA Rogaeve E.I., Sherrington R., Rogaeve E.A., Levesque G., Ikeda M.,  
 RA Liang Y., Chi H., Lin C., Holman K., Tsuda T., Mar L., Sorbi S.,  
 RA Nemias B., Piacentini S., Amaducci L., Chumakov I., Cohen D.,  
 RA Lannfelt L., Fraser P.E., Romenes J.M., St George-Hyslop P.H.;  
 RT "familial Alzheimer's disease in kindreds with missense mutations in  
 RT a gene on chromosome 1 related to the Alzheimer's disease type 3  
 RT gene.";  
 RT Nature 376:775-778(1995).  
 RL  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96109229; PubMed=8618867;  
 RA Li J., Ma J., Potter H.;  
 RT "Identification and expression analysis of a potential familial  
 RT Alzheimer disease gene on chromosome 1 related to AD3.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:12180-12184(1995).  
 RN  
 RP  
 RA SEQUENCE FROM N.A.  
 RA Levy-Lahad E., Poorkaj P., Wang K., Fu Y.H., Oshima J.,  
 RA Mulligan J., Schellenberg G.D.;  
 RT Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RL  
 RN  
 RP SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.  
 RX MEDLINE=96160372; PubMed=8574969;  
 RA Kovacs D.M., Fausett H.J., Page K.J., Kim T.-N., Molr R.D.,  
 RA Merriam D.E., Hollister R.D., Hallmark O.G., Mancini R.,  
 RA Felsenstein K.M., Hyman B.T., Tanzi R.E., Masco W.;  
 RT "Alzheimer-associated presenilin 1 and 2: neuronal expression in  
 RT brain and localization to intracellular membranes in mammalian  
 RT cells.";  
 RL Nat. Med. 2:224-229(1996).  
 RN  
 RP MUTAGENESIS OF ASP-366.  
 RX MEDLINE=99428546; PubMed=10497236;  
 RA Steiner H., Duff K., Capell A., Romig H., Grim M.G., Lincoln S.,  
 RA Hardy J., Yu X., Picciano M., Fechter K., Clifton M., Kopan R.,  
 RA Pesold B., Beck S., Baader M., Tomita T., Iwatsubo T., Baumeister R.,  
 RA Haass C.;  
 RT "A loss of function mutation of presenilin-2 interferes with amyloid  
 RT beta-peptide production and notch signaling.";  
 RL J. Biol. Chem. 274:28669-28673(1999).  
 RN  
 RP  
 RA MUTAGENESIS OF ASP-263 AND ASP-366.  
 RX MEDLINE=20119269; PubMed=10652302;  
 RA Kimberly W.T., Xia W., Rahmati T., Wolfe M.S., Selkoe D.J.;  
 RT "The transmembrane aspartates in presenilin 1 and 2 are obligatory for  
 RT gamma-secretase activity and amyloid beta-protein generation.";  
 RL J. Biol. Chem. 275:3173-3178(2000).  
 RN  
 RP  
 RA REVIEW ON VARIANTS.  
 RX MEDLINE=98180715; PubMed=9521418;  
 RA Cruts M., van Broeckhoven C.;  
 RT "Presenilin mutations in Alzheimer's disease.";  
 RL Hum. Mutat. 11:183-190(1998).  
 RN  
 RP  
 RA VARIANT AD HIS-62.  
 RX MEDLINE=98046005; PubMed=9384602;  
 RA Cruts M., van Duijn C.M., Backhovens H., van den Broeck M.,  
 RA Wehnert A., Serneels S., Sherrington R., Hutton M., Hardy J.,  
 RA St George-Hyslop P.H., Hofman A., van Broeckhoven C.;  
 RT "Estimation of the genetic contribution of presenilin-1 and -2  
 RT mutations in a population-based study of presenile Alzheimer  
 RT disease.";  
 RL Hum. Mol. Genet. 7:43-51(1998).

RN [10]  
 RP VARIANT AD IL6-148.  
 RA Lao J.I., Beyer K., Fernandez-Novoa L., Cacabelos R.;  
 RT "A novel mutation in the predicted TM2 domain of the presenilin 2 gene  
 in Spanish patient with late-onset Alzheimer's disease.";  
 RL Neurogenetics 1:293-296(1998).  
 RN [11]  
 RP VARIANTS EOAD PRO-122 AND IL6-239.  
 RX MEDLINE-20100613; PubMed-10631141;  
 RA Flinch U., Mueller-Thomsen T., Mann U., Eggers C., Marksteiner J.,  
 RA Meins W., Binetti G., Alberici A., Hock C., Nitsch R.M., Gal A.;  
 RT "High prevalence of pathogenic mutations in patients with early-onset  
 dementia detected by sequence analyses of four different genes.";  
 RL Am. J. Hum. Genet. 66:110-117(2000).  
 CC -1- FUNCTION: MAY PLAY A ROLE IN INTRACELLULAR SIGNALING AND GENE  
 EXPRESSION OR IN LINKING CHROMATIN TO THE NUCLEAR MEMBRANE. MAY  
 FUNCTION IN THE CYTOPLASMIC PARTITIONING OF PROTEINS. IS INVOLVED  
 IN THE PROTEOLYTICAL PROCESSING OF AMYLOID PRECURSOR PROTEIN (APP)  
 AND OF NOTCH1.  
 CC -1- SUBUNIT: PREDOMINANTLY HETERODIMER OF A N-TERMINAL (NTF) AND A  
 C-TERMINAL (CTF) ENDOPROTEOLYTIC FRAGMENT. ASSOCIATES WITH  
 NICASTRIN AND WITH PROTEOLYTIC PROCESSED C-TERMINAL FRAGMENTS C83  
 AND C99 OF THE AMYLOID PRECURSOR PROTEIN (APP).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. GOLGI AND  
 ENDOPLASMIC RETICULUM.  
 CC -1- TISSUE SPECIFICITY: THE LARGER FORM IS SEEN IN THE PLACENTA,  
 SKELETAL MUSCLE AND HEART WHILE THE SHORTER FORM IS SEEN IN THE  
 HEART, BRAIN, PLACENTA, LIVER, SKELETAL MUSCLE AND KIDNEY.  
 CC -1- PTM: HETEROGENEOUS PROTEOLYTIC PROCESSING GENERATES N-TERMINAL  
 AND C-TERMINAL FRAGMENTS.  
 CC -1- PTM: PHOSPHORYLATED ON SERINE RESIDUES.  
 CC -1- DISEASE: DEFECTS IN PSEN2 ARE RESPONSIBLE FOR EARLY-ONSET  
 AUTOSOMAL DOMINANT ALZHEIMER DISEASE (EOAD), WHICH IS THE MOST  
 SEVERE FORM OF THE DISEASE, COMPLETE PENETRANCE. AN ONSET  
 OCCURRING AS EARLY AS 30 YEARS OF AGE. THE SECOND FORM IS LATE-  
 ONSET (LOAD) ALSO CALLED THE FAMILIAL ALZHEIMER DISEASE (FAD). AD  
 IS A NEURODEGENERATIVE DISORDER CHARACTERIZED BY PROGRESSIVE  
 DEMENTIA, PARKINSONISM, AND DEPOSITION OF FIBRILLAR AMYLOID  
 PROTEINS AS INTRANEURONAL NEUROFIBRILLARY TANGLES, EXTRACELLULAR  
 AMYLOID PLAQUES AND VASCULAR AMYLOID DEPOSITS.  
 CC -1- DISEASE: THREE CAUSATIVE GENES HAVE BEEN IDENTIFIED THAT WHEN  
 MUTATED LEAD TO PRESENILE ALZHEIMER'S DISEASE: APP (AMYLOID  
 PRECURSOR PROTEIN GENE), PSEN1 AND PSEN2. THESE THREE GENES  
 ACCOUNT FOR HALF OF THE FAMILIES WITH AUTOSOMAL DOMINANT PRESENILE  
 AD, WHICH REPRESENT APPROXIMATELY 10% OF THE WHOLE AD POPULATION.  
 IN ADDITION, APOLOPORTEIN E HAS BEEN IDENTIFIED AS A RISK-  
 MODIFYING LOCUS.  
 CC -1- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: LA3964; AAB59557.1; -;  
 DR EMBL: LA4577; AAC42012.1; -;  
 DR EMBL: U34349; AAC50290.1; -;  
 DR EMBL: U50871; AAB50054.1; -;  
 DR MEROPS: A22.002; -;  
 DR GeneW: HGNC:9509; PSEN2.  
 DR MIM: 600759; -;  
 DR InterPro: IPR001108; Presenilin.  
 DR Pfam: PF01080; Presenilin; 1.  
 DR Trnsmembrane: Phosphorylation; Endoplasmic reticulum; Golgi stack;  
 KM Alzheimer's disease; Disease mutation; Alternative splicing.  
 FT CHAIN 1 297  
 FT CHAIN 298 448  
 FT DOMAIN 1 87  
 FT TRANSMEM 88 108  
 FT POTENTIAL.

FT DOMAIN 109 138  
 FT TRANSMEM 139 159  
 FT DOMAIN 160 166  
 FT TRANSMEM 167 187  
 FT DOMAIN 188 200  
 FT TRANSMEM 201 221  
 FT DOMAIN 222 223  
 FT TRANSMEM 224 244  
 FT DOMAIN 245 249  
 FT TRANSMEM 250 270  
 FT DOMAIN 271 388  
 FT TRANSMEM 389 409  
 FT TRANSMEM 414 434  
 FT VARSPPLIC 263 296  
 FT VARIANT 62 62  
 FT VARIANT 122 122  
 FT VARIANT 141 141  
 FT VARIANT 141 141  
 FT VARIANT 148 148  
 FT VARIANT 239 239  
 FT VARIANT 239 239  
 FT VARIANT 239 239  
 FT MUTAGEN 263 263  
 FT MUTAGEN 366 366  
 FT CONFLICT 123 123  
 FT CONFLICT 325 325  
 FT CONFLICT 358 358  
 FT CONFLICT 432 448  
 SQ SEQUENCE 448 AA; 50140 MM; A927EC623468116 CRC64;  
 Query Match 6.0%; Score 28; DB 1; Length 448;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-19;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 253 ISYVDLVAVLCPRGPRMLVETAOERNE 280  
 |||||||  
 DB 259 ISYVDLVAVLCPRGPRMLVETAOERNE 286  
 RESULT 9  
 PSN2.MOUSE  
 ID PSN2.MOUSE STANDARD; PRT; 448 AA.  
 AC 061144; P97935; P97934; 054977;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Presenilin 2 (PS-2) (Alz-3) (Alzheimer disease 4 homolog).  
 GN PSEN2 OR PSNL2 OR ALG3 OR AD4H OR PS-2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RX TISSUE=Liver;  
 RX MEDLINE-97094660; PubMed-8940094;  
 RX Vito P., Wolozin B., Gangji J.K., Iwasaki K., Iacana E., D'Adamo L.;  
 RT "Requirement of the familial Alzheimer's disease gene PS2 for  
 RT apoptosis. Opposing effect of ALG-3.";  
 RL J. Biol. Chem. 271:31025-31028(1996).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NIH SWISS;  
 RA Sahara N., Mori H., Shirasawa T.;  
 RT "Molecular cloning of mouse presenilin 2 gene.";

RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RC SEQUENCE OF 340-448 FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=96152375; PubMed=8560270;  
 RA Vito P., Lacana E., D'Adamo L.;  
 RT "Interfering with apoptosis: Ca(2+)-binding protein ALG-2 and  
 RT Alzheimer's disease gene ALC-3.";  
 RL Science 271:521-525(1996).  
 CC -1- FUNCTION: MAY PLAY A ROLE IN INTRACELLULAR SIGNALING AND GENE  
 CC EXPRESSION OR IN LINKING CHROMATIN TO THE NUCLEAR MEMBRANE. MAY  
 CC FUNCTION IN THE CYTOPLASMIC PARTITIONING OF PROTEINS (BY  
 CC SIMILARITY).  
 CC -1- SUBUNIT: PREDOMINANTLY HETERODIMER OF A N-TERMINAL (NTF) AND A  
 CC C-TERMINAL (CTF) ENDOPROTEOLYTICAL FRAGMENT (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. GOLGI AND  
 CC ENDOPASMIC RETICULUM (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: UNIDIRECTIONALLY EXPRESSED, MOSTLY IN THE LIVER.  
 CC -1- PTM: PHOSPHORYLATED ON SERINE RESIDUES (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U57324; AAC52937.1; -  
 DR EMBL: U57325; AAC53311.1; -  
 DR EMBL: AF038935; AAB92660.1; -  
 DR EMBL: U49111; AAC52935.1; -  
 DR MEROPS: A22.002; -  
 DR MGD: MGI:109284; Psen2.  
 DR InterPro: IPR001108; Presenilin.  
 DR Pfam: PF01080; Presenilin.1.  
 DR PRINTS: PR01072; PRESENILIN.  
 KM Transmembrane; Phosphorylation; Endoplasmic reticulum; Golgi stack.  
 FT CHAIN 1 297 PRESENILIN 2 NTF SUBUNIT (BY SIMILARITY).  
 FT CHAIN 298 448 PRESENILIN 2 CTF SUBUNIT.  
 FT DOMAIN 1 87 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 88 108 LUMENAL (POTENTIAL).  
 FT DOMAIN 109 138 LUMENAL (POTENTIAL).  
 FT TRANSMEM 139 159 POTENTIAL.  
 FT DOMAIN 160 166 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 167 187 POTENTIAL.  
 FT DOMAIN 188 200 LUMENAL (POTENTIAL).  
 FT TRANSMEM 201 221 POTENTIAL.  
 FT DOMAIN 222 223 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 224 244 POTENTIAL.  
 FT DOMAIN 245 249 LUMENAL (POTENTIAL).  
 FT TRANSMEM 250 270 POTENTIAL.  
 FT DOMAIN 271 388 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 389 409 POTENTIAL.  
 FT TRANSMEM 414 434 POTENTIAL.  
 FT CONFLICT 87 87 R -> V (IN REF. 2).  
 FT CONFLICT 226 226 A -> H (IN REF. 2).  
 FT CONFLICT 324 324 MISSING (IN REF. 2).  
 SQ SEQUENCE 448 AA; 49955 MW; D6ECA9EF6360D411 CRC64;

Query Match 6.0%; Score 28; DB 1; Length 448;  
 Best Local Similarity 100.0%; Pred. No. 4; 1e-19;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AC 088777; 035546; 008947;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Presenilin 2 (PS-2).  
 GN PSEN2 OR PSNL2 OR PS2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Wistar; TISSUE=Brain;  
 RA Frenzel S., Abdel A.S., Luebert H.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Wistar; TISSUE=Brain;  
 RX MEDLINE=97473536; PubMed=9332390;  
 RA Takahashi H., Mercken M., Nakazato Y., Noguchi K., Murayama M.,  
 RA Imahori K., Takashima A.;  
 RT "Cloning of cDNA and expression of the gene encoding rat  
 RT presenilin-2.";  
 RL Gene 197:383-387(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Wistar; TISSUE=Brain;  
 RX MEDLINE=98207716; PubMed=9545577;  
 RA Tanahashi H., Tabira T.;  
 RT "Cloning of the cDNA encoding rat presenilin-2.";  
 RL Biochem. Biophys. Acta 1396:259-262(1998).  
 CC -1- FUNCTION: MAY PLAY A ROLE IN INTRACELLULAR SIGNALING AND GENE  
 CC EXPRESSION OR IN LINKING CHROMATIN TO THE NUCLEAR MEMBRANE. MAY  
 CC FUNCTION IN THE CYTOPLASMIC PARTITIONING OF PROTEINS (BY  
 CC SIMILARITY).  
 CC -1- SUBUNIT: PREDOMINANTLY HETERODIMER OF A N-TERMINAL (NTF) AND A  
 CC C-TERMINAL (CTF) ENDOPROTEOLYTICAL FRAGMENT (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. GOLGI AND  
 CC ENDOPASMIC RETICULUM (BY SIMILARITY).  
 CC -1- PTM: PHOSPHORYLATED ON SERINE RESIDUES (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X99267; CAA67663.1; -  
 DR EMBL: D83700; BAA22832.1; -  
 DR EMBL: AB004454; BAA20406.1; -  
 DR MEROPS: A22.002; -  
 DR InterPro: IPR001108; Presenilin.  
 DR Pfam: PF01080; Presenilin.1.  
 DR PRINTS: PR01072; PRESENILIN.  
 KM Transmembrane; Phosphorylation; Endoplasmic reticulum; Golgi stack.  
 FT CHAIN 1 297 PRESENILIN 2 NTF SUBUNIT (BY SIMILARITY).  
 FT CHAIN 298 448 PRESENILIN 2 CTF SUBUNIT (BY SIMILARITY).  
 FT DOMAIN 1 87 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 88 106 LUMENAL (POTENTIAL).  
 FT DOMAIN 107 141 LUMENAL (POTENTIAL).  
 FT TRANSMEM 142 159 POTENTIAL.  
 FT DOMAIN 160 166 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 167 188 POTENTIAL.  
 FT DOMAIN 189 202 LUMENAL (POTENTIAL).  
 FT TRANSMEM 203 219 POTENTIAL.  
 FT DOMAIN 220 229 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 230 246 POTENTIAL.  
 FT DOMAIN 247 252 LUMENAL (POTENTIAL).  
 FT TRANSMEM 253 269 POTENTIAL.  
 FT DOMAIN 270 386 CYTOPLASMIC (POTENTIAL).



```

FT TRANSMEM 387 406 POTENTIAL.
FT TRANSMEM 414 434 POTENTIAL.
FT DOMAIN 7 7 S -> T (IN REF. 1).
FT CONFLICT 86 87 KH -> ND (IN REF. 3).
SQ SEQUENCE 448 AA; 50051 MW; 299a7c416405046c CRC64;

Query Match
Best Local Similarity 100.0%; Score 28; DB 1; Length 448;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 ISYVDLAVLCPEKGPLRLMVLVTAQERNE 280
DB 259 ISYVDLAVLCPEKGPLRLMVLVTAQERNE 286

RESULT 11
PSN2_BOVIN
ID PSN2_BOVIN STANDARD: PRT; 449 AA.
AC 09XT96;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Presentin 2 (PS-2).
GN PSN2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Sahara N., Shirasawa T., Mori H.;
RT "Molecular cloning of bovine presentin 2 gene."
RT Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: MAY PLAY A ROLE IN INTRACELLULAR SIGNALING AND GENE
CC EXPRESSION OR IN LINKING CHROMATIN TO THE NUCLEAR MEMBRANE. MAY
CC FUNCTION IN THE CYTOPLASMIC PARTITIONING OF PROTEINS (BY
CC SIMILARITY).
CC -I- SUBUNIT: PREDOMINANTLY HETERODIMER OF A N-TERMINAL (NPF) AND A
CC C-TERMINAL (CTF) ENDOPROTEOLYTICAL FRAGMENT (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. GOLGI AND
CC ENDOPLASMIC RETICULUM (BY SIMILARITY).
CC -I- PTM: PHOSPHORYLATED ON SERINE RESIDUES (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE PRESENTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: AF038937; AAD39024.1; -.
DR MEROPS; A22.002; -.
DR InterPro; IPR001108; Presentin.
DR Pfam; PF01080; Presentin; 1.
DR PRINTS; PRO1072; PRESENTIN.
DR Transmembrane; Phosphorylation; Endoplasmic reticulum; Golgi stack.
FT CHAIN 1 298 PRESENTIN 2 NPF SUBUNIT (BY SIMILARITY).
FT CHAIN 299 449 PRESENTIN 2 CTF SUBUNIT (BY SIMILARITY).
FT DOMAIN 1 88 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 89 109 POTENTIAL.
FT TRANSMEM 110 139 LUMENAL (POTENTIAL).
FT TRANSMEM 140 160 POTENTIAL.
FT TRANSMEM 161 167 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 168 188 POTENTIAL.
FT TRANSMEM 189 201 LUMENAL (POTENTIAL).
FT TRANSMEM 202 222 POTENTIAL.
FT TRANSMEM 223 224 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 225 245 POTENTIAL.
FT TRANSMEM 246 250 LUMENAL (POTENTIAL).

```

```

FT TRANSMEM 251 271 POTENTIAL.
FT TRANSMEM 272 389 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 390 410 POTENTIAL.
FT TRANSMEM 415 435 POTENTIAL.
SQ SEQUENCE 449 AA; 50301 MW; 51350D05A02B2A7D CRC64;

Query Match
Best Local Similarity 100.0%; Score 28; DB 1; Length 449;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 ISYVDLAVLCPEKGPLRLMVLVTAQERNE 280
DB 260 ISYVDLAVLCPEKGPLRLMVLVTAQERNE 287

RESULT 12
PSN2_XENLA
ID PSN2_XENLA STANDARD: PRT; 449 AA.
AC 012977;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Presentin beta.
GN PS-BETA.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE-97223465; PubMed-9070286;
RA Tsujimura A., Yasojima K., Hashimoto-Gotoh T.;
RT "Cloning of Xenopus presentin-alpha and -beta cDNAs and their
RT differential expression in oogenesis and embryogenesis."
RT Biochem. Biophys. Res. Commun. 231:392-396(1997).
CC -I- FUNCTION: MAY PLAY A ROLE IN NEGATIVE REGULATION OF APOPTOTIC
CC CASCADES DURING OOGENESIS AND EMBRYOGENESIS. AND IN
CC DEVELOPMENTALLY MATURED TISSUES SUCH AS BRAIN TISSUE.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein (by similarity).
CC -I- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN OVARIES AND TO A LESSER
CC EXTENT IN KIDNEY, BRAIN, EYE AND LUNG. WEAK EXPRESSION IN TESTIS,
CC INTESTINE, LIVER AND HEART. PRESENT IN TRACE AMOUNTS IN SKELETAL
CC MUSCLE.
CC -I- DEVELOPMENTAL STAGE: ABUNDANT IN EARLY STAGES OF OOGENESIS AFTER
CC WHICH IT IS NEARLY CONSTANT.
CC -I- SIMILARITY: BELONGS TO THE PRESENTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: D84428; BAA19571.1; -.
DR MEROPS; A22.002; -.
DR InterPro; IPR001108; Presentin.
DR Pfam; PF01080; Presentin; 1.
DR PRINTS; PRO1072; PRESENTIN.
DR Transmembrane; Glycoprotein.
FT TRANSMEM 92 112 POTENTIAL.
FT TRANSMEM 142 162 POTENTIAL.
FT TRANSMEM 170 190 POTENTIAL.
FT TRANSMEM 204 224 POTENTIAL.
FT TRANSMEM 227 247 POTENTIAL.
FT TRANSMEM 253 273 POTENTIAL.
FT TRANSMEM 291 311 POTENTIAL.
FT TRANSMEM 363 383 POTENTIAL.
FT TRANSMEM 390 410 POTENTIAL.
FT TRANSMEM 415 435 POTENTIAL.

```



FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 449 AA; 50274 MM; 2C2C105F5C73F2B CRC64;  
 Query Match 5.8%; Score 27; DB 1; Length 449;  
 Best Local Similarity 100.0%; Pred. No. 3,9e-18;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 81 HVIMLEVPVTCMVVATIKSVFYT 107  
 ||||||||||||||||||  
 Db 90 HVIMLEVPVTCMVVATIKSVFYT 116  
 RESULT 13  
 PSN\_DROME STANDARD; PRT: 541 AA  
 AC 002194; 002395; 076802; Q9V3L9; Q9TY80; Q9V3S1; 096340;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Presentin homolog (DPS) (DMPs).  
 GN PSN OR PS OR CG18803/CG5868.  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND DEVELOPMENTAL STAGE.  
 RC TISSUE=Embryo; PubMed=9141085;  
 RX MEDLINE=97285868; PubMed=9141085;  
 RA Boulianne G.L., Lyvine-Bar I., Humphreys J.M., Liang Y., Lin C.,  
 RA Rogaeve E., St George-Hyslop P.H.;  
 RT "Cloning and characterization of the Drosophila presentin  
 RT homologue.";  
 RL NeuroReport 8:1025-1029(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGE (SHORT ISOFORM).  
 RC STRAIN=Cañton-S; TISSUE=Embryo, and Head;  
 RX MEDLINE=97260623; PubMed=9106743;  
 RA Hong C.-S., Koo E.H.;  
 RT "Isolation and characterization of Drosophila presentin homolog.";  
 RL NeuroReport 8:665-668(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS), TISSUE SPECIFICITY,  
 RP DEVELOPMENTAL STAGE, SUBCELLULAR LOCATION, AND ALTERNATIVE SPLICING.  
 RX MEDLINE=99279250; PubMed=10349633;  
 RA Ye Y., Fortini M.E.;  
 RT "Characterization of Drosophila presentin and its colocalization with  
 RT Notch during development.";  
 RL Mech. Dev. 79:199-211(1998).  
 RN [4]  
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).  
 RX MEDLINE=99221486; PubMed=10206647;  
 RA Ye Y., Lukinova N., Fortini M.E.;  
 RT "Neurogenic phenotypes and altered Notch processing in Drosophila  
 RT presentin mutants.";  
 RL Nature 398:525-529(1999).  
 RN [5]  
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND DEVELOPMENTAL STAGE.  
 RC STRAIN=Cañton-S;  
 RX MEDLINE=98331525; PubMed=9666900;  
 RA Mafany G., Del-Favero J., Valero R., De Jonghe C., Woodrow S.,  
 RA Hendriks L., Van Broeckhoven C., Gonzalez-Duarte R.;  
 RT "Identification of a Drosophila presentin homologue: evidence of  
 RT alternatively spliced forms.";  
 RL J. Neurogenet. 12:41-54(1998).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99423881; PubMed=10493744;  
 RA Guo Y., Lyvine-Bar I., Zhou L., Boulianne G.L.;  
 RT "Drosophila presentin is required for neuronal differentiation and  
 RT affects notch subcellular localization and signaling.";  
 RL J. Neurosci. 19:8435-8442(1999).

RN [7]  
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA Sutton R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blaise J.R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadiu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jaisai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclele J.M.,  
 RA Palazotto M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 CC -I- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE PATTERNING OF THE  
 CC OPTIC LOBES. PROTEIN MAY ACT, TOGETHER WITH NOTCH, TO SPECIFY  
 CC CELL FATES THROUGHOUT DEVELOPMENT.  
 CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. GOLGI AND  
 CC ENDOPLASMIC RETICULUM (BY SIMILARITY).  
 CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A  
 CC SHORT FORM/PS-D; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -I- TISSUE SPECIFICITY: MATERNALLY EXPRESSED IN NURSE AND POLICLE  
 CC CELLS. IN EARLY EMBRYOS, EXPRESSED IN ALL OR MOST CELLS AND LATER  
 CC INCREASES IN CNS AND EPIDERMAL TISSUES. IN LARVAE, EXPRESSION IS  
 CC SEEN IN ALL IMAGINAL DISKS, BRAIN AND OPTIC LOBES. IN PUPAE,  
 CC EXPRESSION IS SEEN IN EYE DISK AND OPTIC LOBES.  
 CC -I- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY  
 CC THROUGHOUT DEVELOPMENT.  
 CC -I- SIMILARITY: BELONGS TO THE PRESENTIN FAMILY.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: U77934; AAB61139.1; -;  
 CC EMBL: U78084; AAB53369.1; -;  
 CC EMBL: AF084184; AAC33129.1; -;  
 CC EMBL: AF084184; AAC33128.1; -;  
 CC EMBL: AF017024; AAD01610.1; -;  
 CC EMBL: AF017025; AAD01611.1; -;  
 CC EMBL: AF017026; AAD01612.1; -;

```

DR EMBL; AF093402; AAD52707.1; -.
DR EMBL; AF093402; AAD52708.1; -.
DR EMBL; AE003591; -. NOT ANNOTATED_CDS.
DR EMBL; AE003591; AAF51598.1; -.
DR MEROPS; A22.0PW; -.
DR FlyBase; FBgn0019947; Psn.
DR InterPro; IPR001108; Presenilin.
DR Pfam; PF01080; Presenilin; 1.
DR PRINTS; PR01072; PRESENILIN.
KM Transmembrane; Glycoprotein; Alternative splicing.
FT TRANSMEM 107 127
FT TRANSMEM 155 175
FT TRANSMEM 183 203
FT TRANSMEM 217 237
FT TRANSMEM 243 263
FT TRANSMEM 266 286
FT TRANSMEM 304 324
FT TRANSMEM 482 502
FT TRANSMEM 507 527
FT CARBOHYD 129 129
FT CARBOHYD 339 339
FT CARBOHYD 410 410
FT VARSPIC 384 397
FT CONFLICT 80 81
SQ SEQUENCE 541 AA; 59304 MW; A3B3D54348A2C03F CRC64;

Query Match 4.7%; Score 22; DB 1; Length 541;
Best Local Similarity 100.0%; Pred. No. 3; le-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 406 TTACFAVAILIGLCTLLAI 427
Db 480 TTACFAVAILIGLCTLLAI 501

RESULT 14
PSN_CAEEL STANDARD; PRT; 444 AA.
AC P52166; Q20076; Q9U9C7;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Presenilin sel-12.
GN SEL-12 OR F35H12.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditia; Rhabditioidea;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A. AND VARIANT SER-60.
RC STRAIN-Bristol N2;
RX MEDLINE=96032531; PubMed=7566091;
RA Levitan D., Greenwald I.;
RT "Facilitation of lin-12-mediated signalling by sel-12, a
RL Caenorhabditis elegans S182 Alzheimer's disease gene.";
RN Nature 377:351-354(1995).
RN [2]
RP REVISIONS TO 84-85.
RA Levitan D.;
RN Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC STRAIN-Bristol N2;
RX MEDLINE=20372200; PubMed=10917532;
RA Wittenburg N., Eimer S., Lakowski B., Roehrig S., Rudolph C.,
BA Baumeister R.;
RT "Presenilin is required for proper morphology and function of neurons
in C. elegans.";
RN Nature 406:306-309(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
OX Nelson J., Gattung S.;

```

```

RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [5]
RP REVISIONS.
RA Materston R.;
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY FACILITATE LIN-12 MEDIATED RECEPTION OF
CC INTERCELLULAR SIGNALS. IT MIGHT BE DIRECTLY INVOLVED IN LIN-12
CC MEDIATED RECEPTION, FUNCTIONING AS A CO-RECEPTOR OR AS A
CC DOWNSTREAM EFFECTOR THAT IS ACTIVATED UPON LIN-12 ACTIVATION.
CC ALTERNATIVELY IT MAY BE INVOLVED IN A MORE GENERAL CELLULAR
CC PROCESS SUCH AS RECEPTOR LOCALIZATION OR RECYCLING AND HENCE
CC INFLUENCE LIN-12 ACTIVITY INDIRECTLY. Required cell-autonomously
CC for correct neurite connectivity of the AIX cholinergic
CC interneurons and their correct functioning in thermotaxis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed in most neurons.
CC -!- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U35660; AAA85511.1; -.
DR EMBL; AF171064; AAD50991.1; -.
DR EMBL; U41540; AAK39230.1; -.
DR MEROPS; A22.0PW; -.
DR WormPep; F35H12.3; CE24946.
DR InterPro; IPR001686; Nema-presenilin.
DR InterPro; IPR001108; Presenilin.
DR Pfam; PF01080; Presenilin; 1.
DR PRINTS; PR01072; PRESENILIN.
DR PRINTS; PR01075; PRESENILINSEL.
KM Transmembrane.
FT TRANSMEM 46 66
FT TRANSMEM 102 122
FT TRANSMEM 131 151
FT TRANSMEM 164 184
FT TRANSMEM 190 210
FT TRANSMEM 213 233
FT TRANSMEM 251 271
FT TRANSMEM 385 405
FT VARIANT 60 60
FT CONFLICT 413 444
SQ SEQUENCE 444 AA; 50034 MW; 37ADB124E16429C CRC64;

Query Match 3.9%; Score 18; DB 1; Length 444;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 EERGKVLGLDGFIFYSVL 392
Db 354 EERGKVLGLDGFIFYSVL 371

RESULT 15
PSNH_ARATH STANDARD; PRT; 453 AA.
ID PSNH_ARATH
AC 064668;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Presenilin homolog.
GN AT1G08700 OR F22013.18.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

```

```

OX NCBI_TaxID=3702:
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia:
RA MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Brehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Malti R., Marzilli A.,
RA Maltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.T.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RL Nature 408:816-820(2000).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PRESENTILIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AC003981; AAF9776.1; -.
DR MEROPS: A2.2.UPW; -.
DR InterPro: IPR001108; Presentilin.
DR Pfam: PF01080; Presentilin.1.
DR PRINTS: PRO1072; PRESENTILIN.
KW Transmembrane.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 69 89 POTENTIAL.
FT TRANSMEM 104 124 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 160 180 POTENTIAL.
FT TRANSMEM 181 201 POTENTIAL.
FT TRANSMEM 370 390 POTENTIAL.
FT TRANSMEM 393 413 POTENTIAL.
FT TRANSMEM 418 438 POTENTIAL.
SQ SEQUENCE 453 AA; 49308 MW; 7FBDD9BB6D97955E CRC64;

Query Match 3.2%; Score 15; DB 1; Length 453;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 KLGSDPFYSVLVG 394
Db 369 KLGSDPFYSVLVG 383

RESULT 16
YGBQ_ECOLI STANDARD: PRT: 103 AA.
AC 046894;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ygbQ.
GN YGBQ OR B2748 OR Z4056 OR ECS3602.
OS Escherichia coli, and
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

```

```

OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RA MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Brluch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor M., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U29579; AAA69258.1; -.
DR EMBL: AE000358; AAC75790.1; -.
DR EMBL: AE005502; AAG57855.1; -.
DR EMBL: AP002562; BAB37025.1; -.
DR EcoGene: EG3111; ygbQ.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 103 AA; 11622 MW; 9B50A7EF637809D7 CRC64;

Query Match 1.7%; Score 8; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 420 LTLLELAI 427
Db 4 LTLLELAI 11

RESULT 17
22P1_RAT STANDARD: PRT: 176 AA.
AC P22282; O63674;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cystatin related protein 1 precursor (CRP-1) (Prostatic 22 kDa
DE glycoprotein P22K16/P22K20) (Androgen regulated 20 kDa protein).
GN CRP1.
OS Rattus norvegicus (Rat).

```



```

CC -----
DR EMBL: U76207; AAC53028.1; -
DR EMBL: Y07621; CAA68900.1; -
DR EMBL: AF303001; AAC40769.1; -
DR MGD: MGI:109619; Neurog2.
DR InterPro: IPR001092; HLH_basic.
DR Pfam: PF00010; HLH; 1.
DR SMART: SMO0353; HLH; 1.
DR PROSITE: PS00038; HLH_1; FALSE_NEG.
DR DNA-binding: Nuclear protein; Transcription regulation; Activator;
KW Neurogenesis; Developmental protein; Differentiation.
FT DNA_BIND 113 124 BASIC DOMAIN.
FT DOMAIN 125 165
FT CONFLICT 60 60 E->G (IN REF. 2).
SQ SEQUENCE 263 AA; 28215 MW; 817EF8246BDCABE CRC64;

Query Match 1.7%; Score 8; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 66 DEEDEEL 73
   |||||
Db 38 DEEDEEL 45

RESULT 19
DCOR_LEIDO STANDARD; PRT; 707 AA.
ID DCOR_LEIDO
AC P27116;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Ornithine decarboxylase (EC 4.1.1.17) (ODC).
OS Leishmania donovani.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
NCBI_TaxID=5661;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92129311; Pubmed=1339439;
RA Hanson S.S., Adelman J., Ullman B.;
  "Amplification and molecular cloning of the ornithine decarboxylase
  gene of Leishmania donovani."
RL J. Biol. Chem. 267:2350-2359(1992).
CC -1- CATALYTIC ACTIVITY: L-ornithine - putrescine + CO(2).
CC -1- Cofactor: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: FIRST STEP AND ALSO THE RATE-LIMITING STEP IN THE PATHWAY
  OF POLYAMINE BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGinine
  DECARBOXYLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M81192; AAA29259.1; -
DR PIR: A42322; A42322.
DR PIR: S27821; S27821.
DR HSP: P00860; 70DC.
DR InterPro: IPR000183; Decarboxylase2.
DR Pfam: PF00278; Orn_DAP_Arg_dec; 1.
DR Pfam: PF02784; Orn_Arg_dec_N; 1.
DR PRINTS: PR01179; ODAICRBLASE.
DR PROSITE: PS00878; ODR_DC_2.1; 1.
DR PROSITE: PS00879; ODR_DC_2.2; 1.
DR Lyase: Decarboxylase; Pyridoxal phosphate; Polyamine biosynthesis.
FT BINDING 288 288 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT ACT_SITE 634 634 BY SIMILARITY.
SQ SEQUENCE 707 AA; 77396 MW; EBDFF14F791EC572D CRC64;

```

```

Query Match 1.7%; Score 8; DB 1; Length 707;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 184 EVFKTYNV 191
   |||||
Db 168 EVFKTYNV 175

RESULT 20
MDRL_CRIGR STANDARD; PRT; 1276 AA.
ID MDRL_CRIGR
AC P21448;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Multidrug resistance protein 1 (P-glycoprotein 1).
GN ABCB1 OR PGY1 OR PGP1.
OS ABCB1 OR griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OC NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92135896; Pubmed=1685679;
RA Endicott J.A., Sarangi F., Ling V.;
  "Complete cDNA sequences encoding the Chinese hamster P-glycoprotein
  gene family."
RL DNA Seq. 2:89-104(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91154265; Pubmed=1671863;
RA Devine S.E., Hussain A., Davide J.P., Melera P.W.;
  "Full length and alternatively spliced pgp1 transcripts in multidrug-
  resistant Chinese hamster lung cells."
RL J. Biol. Chem. 266:4545-4555(1991).
RN [3]
RP SEQUENCE OF 706-1276 FROM N.A.
RX MEDLINE=88122132; Pubmed=2893255;
RA Endicott J.A., Juranka P.F., Sarangi F., Gerlach J.H., Deuchars K.L.,
  Ling V.;
  "Simultaneous expression of two P-glycoprotein genes in
  drug-sensitive Chinese hamster ovary cells."
RL Mol. Cell. Biol. 7:4075-4081(1987).
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
  DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: PGP1 AND PGP2 CAN MEDIATE MDR, WHILE PGP3 APPARENTLY
  CANNOT.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M60040; AAA68883.1; -
DR EMBL: M59253; AAA37004.1; -
DR EMBL: M17897; AAA37006.1; -
DR PIR: A38696; DVHY1C.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR001440; ABCtransportr.
DR Pfam: PF00005; ABC_tran; 2.
DR Pfam: PF00664; ABC_membrane; 2.
DR PRODOM: PD00006; ABC_transportr; 2.
DR SMART: SMO0382; AAA; 2.

```

```
DR PROSITE; PS00211: ABC_TRANSPORTER; 2.
KW APP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KM Multigene family.
FT DOMAIN 1 50 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 117 71 POTENTIAL.
FT TRANSMEM 186 137 POTENTIAL.
FT TRANSMEM 206 206 POTENTIAL.
FT TRANSMEM 213 233 POTENTIAL.
FT TRANSMEM 294 314 POTENTIAL.
FT TRANSMEM 323 343 POTENTIAL.
FT TRANSMEM 344 707 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 708 728 POTENTIAL.
FT TRANSMEM 754 774 POTENTIAL.
FT TRANSMEM 830 850 POTENTIAL.
FT TRANSMEM 851 871 POTENTIAL.
FT TRANSMEM 934 954 POTENTIAL.
FT TRANSMEM 971 991 POTENTIAL.
FT DOMAIN 992 1276 POTENTIAL.
FT NP_BIND 424 431 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 1067 1074 ATP (POTENTIAL).
FT REPEAT 1 635 ATP (POTENTIAL).
FT REPEAT 636 1276
FT CONFLICT 338 339
SQ SEQUENCE 1276 AA; 140925 MW; 44F3F92A186B4DF6 CRC64;

Query Match 1.7%; Score 8; DB 1; Length 1276;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 420 LTLLLLAI 427
| 11111111
Db 854 LTLLLLAI 861

RESULT 21
MDR3_MOUSE STANDARD; PRT; 1276 AA.
ID MDR3_MOUSE
AC P21447;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE Multidrug resistance protein 3 (P-glycoprotein 3) (MDR1A).
GN ABCB1A OR ABCB4 OR PGY3 OR PGY-3 OR MDR3 OR MDR1A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90205845; PubMed=1969610;
RA Devault A., Gros P.;
RT "Two members of the mouse mdr gene family confer multidrug resistance
RT with overlapping but distinct drug specificities."
RL Mol. Cell. Biol. 10:1652-1663(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90287150; PubMed=1972547;
RA Hsu S.I.H., Cohen D., Kirschner L.S., Lothstein L., Hattstein M.,
RA Horwitz S.B.;
RT "Structural analysis of the mouse mdr1a (P-glycoprotein) promoter
RT reveals the basis for differential transcript heterogeneity in
RT multidrug-resistant J774.2 cells."
RL Mol. Cell. Biol. 10:3596-3606(1990).
RN [3]
RP SEQUENCE OF 173-1276 FROM N.A.
RC STRAIN=BA1B/C;
RX MEDLINE=89308614; PubMed=2473069;
RA Hsu S.I.H., Lothstein L., Horwitz S.B.;
RT "Differential overexpression of three mdr gene family members in
RT multidrug-resistant J774.2 mouse cells. Evidence that distinct P-
RT glycoprotein precursors are encoded by unique mdr genes."
RL J. Biol. Chem. 264:12053-12062(1989).
RC - FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
```

```
CC CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC - SUBCELLULAR LOCATION: Integral membrane protein.
CC - MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE
CC RELATED BUT DISTINCT CELLULAR GENES.
CC - SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M30697; AAA39517.1; -
DR EMBL: M33581; AAA39514.1; -
DR EMBL: M33580; AAA39518.1; -
DR EMBL: M24417; AAA03243.1; -
DR PIR: A34175; DVMSIA.
DR PIR: A34786; A34786.
DR MGI: MGI:97570; Abcb1a.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; ABC_transport.
DR InterPro: IPR001140; ABCtransportTM.
DR Pfam: PF00005; ABC_tran; 2.
DR Pfam: PF00664; ABC_membrane; 2.
DR ProDom: PD000006; ABC_transport; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE; PS00211: ABC_TRANSPORTER; 2.
KW APP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KM Multigene family.
FT DOMAIN 1 47 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 48 71 POTENTIAL.
FT TRANSMEM 117 136 POTENTIAL.
FT TRANSMEM 186 205 POTENTIAL.
FT TRANSMEM 212 232 POTENTIAL.
FT TRANSMEM 298 318 POTENTIAL.
FT TRANSMEM 327 346 POTENTIAL.
FT TRANSMEM 347 707 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 708 728 POTENTIAL.
FT TRANSMEM 753 773 POTENTIAL.
FT TRANSMEM 829 848 POTENTIAL.
FT TRANSMEM 853 880 POTENTIAL.
FT TRANSMEM 942 961 POTENTIAL.
FT TRANSMEM 964 984 POTENTIAL.
FT TRANSMEM 985 1276 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 423 430 ATP (BY SIMILARITY).
FT NP_BIND 1066 1073 ATP (BY SIMILARITY).
FT REPEAT 1 632
FT REPEAT 633 1276
FT CARBOHYD 83 83
FT CARBOHYD 87 87
FT CARBOHYD 90 90
FT CONFLICT 526 527 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 939 939 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 939 939 OL -> HV (IN REF. 2 AND 3).
FT CONFLICT 1036 1036 F -> S (IN REF. 2 AND 3).
FT CONFLICT 1036 1036 F -> V (IN REF. 2 AND 3).
SQ SEQUENCE 1276 AA; 140754 MW; 75C71F3BF1F58481 CRC64;

Query Match 1.7%; Score 8; DB 1; Length 1276;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 420 LTLLLLAI 427
| 11111111
Db 853 LTLLLLAI 860

RESULT 22
MDRL_HUMAN STANDARD; PRT; 1280 AA.
ID MDRL_HUMAN
AC P08183; Q12755; Q14812;
DT 01-NOV-1988 (Rel. 08, Created)
DT 01-MAY-1997 (Rel. 35, Last sequence update)
```





```
FT CONFLICT 185 185 G -> V (IN REF. 1 AND 3).
FT CONFLICT 336 336 MISSING (IN REF. 3).
FT CONFLICT 412 412 G -> A (IN REF. 3).
FT CONFLICT 438 438 0 -> S (IN REF. 3).
SQ SEQUENCE 1280 AA; 141462 MW; AB1C279531F43675 CRC64;

Query Match
Best Local Similarity 1.78; Score 8; DB 1; Length 1280;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 420 LLLLLL 427
DB 857 LLLLLL 864

RESULT 23
TEMP_RANTE
ID TEMP_RANTE STANDARD; PRT; 58 AA.
AC P79876; P56922;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Temporin H precursor.
OS Rana temporaria (European common frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=8407;
RN SEQUENCE FROM N.A., SEQUENCE OF 47-56, AND SYNTHESIS OF 47-56.
RC TISSUE=Skin.
RX MEDLINE=97175050; PubMed=9022710;
RA Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,
RA Baria D.;
RT "Temporins, antimicrobial peptides from the European red frog Rana
RT temporaria.";
RL Eur. J. Biochem. 242:788-792(1996).
CC -1- FUNCTION: HAS NO ANTIBACTERIAL ACTIVITY.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y09394; CAA70563.1; -.
CC DR InterPro: IPR004275; Brevinin.
CC DR Pfam: PF03032; Brevinin.1.
CC KW Amphibian skin; Amidation; Signal; Multigene family.
CC FT SIGNAL 1 22 POTENTIAL.
CC FT PROPEP 23 44
CC FT CHAIN 47 56 TEMPORIN H.
CC FT MOD_RES 56 56 AMIDATION (G-57 PROVIDE AMIDE GROUP).
CC SQ SEQUENCE 58 AA; 6806 MW; 684AECB0451E5E81 CRC64;

Query Match
Best Local Similarity 1.58; Score 7; DB 1; Length 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SLLILFF 176
DB 7 SLLILFF 13

RESULT 24
TEMP_RANTE
ID TEMP_RANTE STANDARD; PRT; 61 AA.
AC P79874;
DT 01-NOV-1997 (Rel. 35, Created)
```

```
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Temporin B precursor.
OS Rana temporaria (European common frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=8407;
RN SEQUENCE FROM N.A., SEQUENCE OF 47-59, AND SYNTHESIS OF 47-59.
RC TISSUE=Skin.
RX MEDLINE=97175050; PubMed=9022710;
RA Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,
RA Baria D.;
RT "Temporins, antimicrobial peptides from the European red frog Rana
RT temporaria.";
RL Eur. J. Biochem. 242:788-792(1996).
CC -1- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM-POSITIVE
CC BACTERIA.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y09393; CAA70562.1; -.
CC DR InterPro: IPR004275; Brevinin.
CC DR Pfam: PF03032; Brevinin.1.
CC KW Amphibian skin; Antibiotic; Amidation; Signal; Multigene family.
CC FT SIGNAL 1 22 POTENTIAL.
CC FT PROPEP 23 44
CC FT CHAIN 47 59 TEMPORIN B.
CC FT MOD_RES 59 59 AMIDATION (G-60 PROVIDE AMIDE GROUP).
CC SQ SEQUENCE 61 AA; 7101 MW; B73F75689C300357 CRC64;

Query Match
Best Local Similarity 1.58; Score 7; DB 1; Length 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SLLILFF 176
DB 7 SLLILFF 13

RESULT 25
TEMP_RANTE
ID TEMP_RANTE STANDARD; PRT; 61 AA.
AC P79875;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Temporin G precursor.
OS Rana temporaria (European common frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=8407;
RN SEQUENCE FROM N.A.
RN [1]
RC TISSUE=Skin.
RX MEDLINE=97175050; PubMed=9022710;
RA Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,
RA Baria D.;
RT "Temporins, antimicrobial peptides from the European red frog Rana
RT temporaria.";
RL Eur. J. Biochem. 242:788-792(1996).
CC -1- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM-NEGATIVE AND
CC GRAM-POSITIVE BACTERIA.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -----
```



```

CC -1 SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y09395; CAA70564.1; -.
CC InterPro: IPR004275; Brevinin.1.
CC Pfam: PF03032; Brevinin.1.
CC KW Amphibian skin; Antibiotic; Amidation; Signal; Multigene family.
CC SIGNAL 1 22 POTENTIAL.
CC PROPEP 23 44 POTENTIAL.
CC CHAIN 47 59 TEMPORIN G.
CC MOD_RES 59 59 AMIDATION (G-60 PROVIDE AMIDE GROUP).
CC SEQUENCE 61 AA: 7171 MW: 8458.670 Da EDFA8BC79DFD9F2 CRC64:
CC
CC Query Match 1.5%; Score 7; DB 1; Length 61;
CC Best Local Similarity 100.0%; Pred. No. 16;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC Oy 170 SLLLF 176
CC | | | | |
CC Db 7 SLLLF 13
CC
CC RESULT 26
CC GAES_RANRU STANDARD: PRT: 65 AA.
CC ID P80399; Q81329;
CC DT 01-NOV-1995 (Rel. 32, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Gaegurin-5 precursor.
CC OS Rana rugosa (Wrinkled frog).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
CC NCBI_TaxID=8410;
CC [1]
CC RP SEQUENCE FROM N.A.
CC RC Tissue-Skin;
CC RA Park J.M., Lee J.Y., Moon H.M., Lee B.J.;
CC RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC RN [2]
CC RP SEQUENCE OF 42-65.
CC RC Tissue-Skin;
CC RX MEDLINE=95091844; PubMed=7999137;
CC RA Park J.M., Jung J.-E., Lee B.J.;
CC RT "Antimicrobial peptides from the skin of a Korean frog, Rana
CC rugosa."
CC RL Biochem. Biophys. Res. Commun. 205:948-954(1994).
CC -1- FUNCTION: HAS A NON-HEMOCLYTIC ACTIVITY. HAS A BROAD SPECTRUM OF
CC ACTIVITY AGAINST BOTH GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA,
CC FUNGI AND PROTOZOA.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: SKIN.
CC -1 SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U22393; AAA64412.1; -.

```

```

DR InterPro: IPR004275; Brevinin.
DR Pfam: PF03032; Brevinin.1.
DR KW Amphibian skin; Antibiotic; Signal.
DR SIGNAL 1 22 POTENTIAL.
DR PROPEP 23 39 POTENTIAL.
DR CHAIN 42 65 GAEGURIN-5.
DR DISULFID 59 65 BY SIMILARITY.
DR CONFLICT 58 58 F->K (IN REF. 2).
DR SEQUENCE 65 AA: 7414 MW: 8458.670 Da EDFA8BC79DFD9F2 CRC64:
DR
DR Query Match 1.5%; Score 7; DB 1; Length 65;
DR Best Local Similarity 100.0%; Pred. No. 17;
DR Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DR
DR Oy 170 SLLLF 176
DR | | | | |
DR Db 7 SLLLF 13
DR
DR RESULT 27
DR RLXN_RANCA STANDARD: PRT: 66 AA.
DR ID P39084;
DR DT 01-FEB-1995 (Rel. 31, Created)
DR DT 01-FEB-1995 (Rel. 31, Last sequence update)
DR DT 15-JUL-1998 (Rel. 36, Last annotation update)
DR DE Ranalexin precursor.
DR OS Rana catesbeiana (Bull frog).
DR OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DR Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
DR NCBI_TaxID=8400;
DR [1]
DR RP SEQUENCE FROM N.A., AND SEQUENCE OF 47-66.
DR RX MEDLINE=94193792; PubMed=8144672;
DR RA Clark D.P., Durrell S., Maloy W.L., Zasloff M.;
DR RT "Ranalexin. A novel antimicrobial peptide from bullfrog (Rana
DR catesbeiana) skin, structurally related to the bacterial antibiotic,
DR polymyxin."
DR RL J. Biol. Chem. 269:10849-10855(1994).
DR [2]
DR RP STRUCTURE BY NMR OF RANALEXIN.
DR RX MEDLINE=98237592; PubMed=9578480;
DR RA Vignal E., Chavanteau A., Roch P., Chiche L., Grassy G., Calas B.,
DR Annelas A.;
DR RT "Solution structure of the antimicrobial peptide ranalexin and a
DR study of its interaction with perdeuterated dodecylphosphocholine
DR micelles."
DR RL Eur. J. Biochem. 253:221-228(1998).
DR -1- FUNCTION: POTENT MICROBICIDAL ACTIVITY, ACTIVE AGAINST
DR S.AUREUS AND E.COLI. IT ACTS AS WELL AS A MEMBRANE-DISRUPTIVE
DR AGENT AT HIGHER CONCENTRATIONS.
DR -1- SUBCELLULAR LOCATION: Secreted.
DR -1- TISSUE SPECIFICITY: GRANULAR GLANDS OF SKIN.
DR -1- DEVELOPMENTAL STAGE: EXPRESSION STARTS AT METAMORPHOSIS AND
DR CONTINUES INTO ADULTHOOD.
DR -1 SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
DR FAMILY.
DR -----
DR This SWISS-PROT entry is copyright. It is produced through a collaboration
DR between the Swiss Institute of Bioinformatics and the EMBL outstation -
DR the European Bioinformatics Institute. There are no restrictions on its
DR use by non-profit institutions as long as its content is in no way
DR modified and this statement is not removed. Usage by and for commercial
DR entities requires a license agreement (See http://www.isb-sib.ch/announce/
DR or send an email to license@isb-sib.ch).
DR -----
DR EMBL: S69903; AAB30394.1; -.
DR PIR: A53744; A53744.
DR InterPro: IPR004275; Brevinin.
DR Pfam: PF03032; Brevinin.1.
DR KW Amphibian skin; Antibiotic; Hemolysis; Signal.
DR SIGNAL 1 20 POTENTIAL.
DR PROPEP 21 44 SMALL ACIDIC PEPTIDE.

```

FT PEPTIDE 47 66 RANALEXIN.  
DT DISULFID 60 66  
SQ SEQUENCE 66 AA: 7615 MW: 096B8AD58A3C8513 CRC64;  
Query Match 1.5%; Score 7; DB 1; Length 66;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 170 SLLLF 176  
DB 7 SLLLF 13  
RESULT 28  
ATPL\_BACCA STANDARD; PRT; 72 AA.  
ID ATPL\_BACCA PA1015;  
AC PA1015;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE ATP synthase C chain (EC 3.6.3.14) (Lipid-binding protein).  
GN ATP.  
OS Bacillus caldotenax.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1395;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Ishizuka M.;  
RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: THIS IS ONE OF THE THREE CHAINS OF THE NONENZYMATIC  
CC COMPONENT (CF(0) SUBUNIT) OF THE ATPASE COMPLEX.  
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(in) = ADP + phosphate +  
CC H(+)(out).  
CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC  
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE  
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)  
CC HAS THREE MAIN SUBUNITS: A, B AND C.  
CC -1- MISCELLANEOUS: DICICLOHEXYLCARBODIIMIDE (DCCD) INHIBITS ATPASE.  
CC -1- SIMILARITY: BELONGS TO THE ATPASE C CHAIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC  
DR EMBL: D38057; BAA07244.1; -.  
DR HSSP: P00844; 1A91.  
DR InterPro: IPR002379; ATPase\_Csub.  
DR InterPro: IPR000454; Eub\_ATPase\_Csub.  
DR Pfam: PF00137; ATP-synt\_C; 1.  
DR PRINTS: PR00124; ATPASEC.  
DR TIGRFAMS: TIGR01260; ATP\_synt\_c; 1.  
DR TRSFAMS: PS00605; ATPASE\_C; FALSE\_NEG.  
KW Hydrogen ion transport; CF(0); Transmembrane; Lipid-binding.  
FT BINDING 56 56 DICICLOHEXYLCARBODIIMIDE (BY SIMILARITY).  
SQ SEQUENCE 72 AA: 7249 MW: 211ED1F2A538B531 CRC64;  
Query Match 1.5%; Score 7; DB 1; Length 72;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 177 FSFTYIG 183  
DB 65 FSFTYIG 71  
RESULT 29  
ATPL\_BACP3 STANDARD; PRT; 72 AA.  
ID ATPL\_BACP3  
OC P00845;  
DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE ATP synthase C chain (EC 3.6.3.14) (Lipid-binding protein).  
GN ATP.  
OS Bacillus PS3 (Thermophilic bacterium PS-3).  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=70306;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88163679; PubMed=2894854;  
RA Ohta S., Yohda M., Ishizuka M., Hirata H., Hamamoto T.,  
RA Otawara-Hamamoto Y., Matsuda K., Kagawa Y.;  
RT "Sequence and over-expression of subunits of adenosine triphosphate  
RT synthase in thermophilic bacterium PS3.";  
RL Biochim. Biophys. Acta 933:141-155(1988).  
RN [2]  
RP SEQUENCE.  
RX MEDLINE=80245993; PubMed=6447066;  
RA Hoppe J., Sebald W.;  
RT "Amino acid sequence of the proteolipid subunit of the proton-  
RT translocating ATPase complex from the thermophilic bacterium PS-3.";  
RL Eur. J. Biochem. 107:57-65(1980).  
CC -1- FUNCTION: THIS IS ONE OF THE THREE CHAINS OF THE NONENZYMATIC  
CC COMPONENT (CF(0) SUBUNIT) OF THE ATPASE COMPLEX.  
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(in) = ADP + phosphate +  
CC H(+)(out).  
CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC  
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE  
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)  
CC HAS THREE MAIN SUBUNITS: A, B AND C.  
CC -1- MISCELLANEOUS: DICICLOHEXYLCARBODIIMIDE (DCCD) INHIBITS ATPASE.  
CC -1- SIMILARITY: BELONGS TO THE ATPASE C CHAIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC  
DR EMBL: X07804; CAA30649.1; -.  
DR PIR: A01048; LMHWA3.  
DR PIR: S02255; S02255.  
DR HSSP: P00844; 1A91.  
DR InterPro: IPR002379; ATPase\_Csub.  
DR InterPro: IPR000454; Eub\_ATPase\_Csub.  
DR Pfam: PF00137; ATP-synt\_C; 1.  
DR PRINTS: PR00124; ATPASEC.  
DR TIGRFAMS: TIGR01260; ATP\_synt\_c; 1.  
DR PROSITE: PS00605; ATPASE\_C; 1.  
KW Hydrogen ion transport; CF(0); Transmembrane; Lipid-binding;  
FT FORMYLATION 1 1  
FT MOD\_RES 56 56 DICICLOHEXYLCARBODIIMIDE (BY SIMILARITY).  
SQ SEQUENCE 72 AA: 7334 MW: 211ED1F2BD740FE1 CRC64;  
Query Match 1.5%; Score 7; DB 1; Length 72;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 177 FSFTYIG 183  
DB 65 FSFTYIG 71  
RESULT 30  
ATPL\_BACST STANDARD; PRT; 72 AA.  
ID ATPL\_BACST  
AC P42011;  
DT 01-NOV-1995 (Rel. 32, Created)

DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 RA ATP synthase C chain (EC 3.6.3.14) (Lipid-binding protein).  
 GN ATP.  
 OS Bacillus stearothermophilus.  
 CC Bacteria; Firmicutes; Bacillales; Geobacillus.  
 NC NCB1\_TaxID=1422;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ishizuka M., Kamet T.;  
 RL Submitted (AUG-1994) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: THIS IS ONE OF THE THREE CHAINS OF THE NONENZYMATIC  
 CC COMPONENT (CF(0) SUBUNIT) OF THE ATPASE COMPLEX.  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(in) = ADP + phosphate +  
 CC H(+)(out).  
 CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC  
 CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE  
 CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)  
 CC HAS THREE MAIN SUBUNITS: A, B AND C.  
 CC -1- MISCELLANEOUS: DICYCLOHEXYLCARBODIIMIDE (DCCD) INHIBITS ATPASE.  
 CC -1- SIMILARITY: BELONGS TO THE ATPASE C CHAIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D38059; BAA07251.1; -;  
 DR HSSP: P00844; I491.  
 DR InterPro: IPR002379; ATPase\_Csub.  
 DR InterPro: IPR000454; Eub\_ATPase\_Csub.  
 DR Pfam: PF00137; ATP\_synC\_1.  
 DR PRINTS: PR00124; ATPSEC.  
 DR TIGRFAMs: TIGR01260; ATP\_synC\_1.  
 DR PROSITE: PS00605; ATPase\_C\_1.  
 KW Hydrogen ion transport; CF(0); Transmembrane; Lipid-binding.  
 FT BINDING 56 56 DICYCLOHEXYLCARBODIIMIDE (BY SIMILARITY).  
 SO SEQUENCE 72 AA: 7348 MW: 211ED564661D2FE1 CRC64;  
 Query Match 1.5%; Score 7; DB 1; Length 72;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 177 FSEYILG 183  
 Db 65 FSEYILG 71  
 RESULT 31  
 YMJ0\_CAEEL STANDARD: PRT: 106 AA.  
 ID YMJ0\_CAEEL  
 AC P34485;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein F59B2.10 in Chromosome III.  
 GN F59B2.10.  
 OS Caenorhabditis elegans.  
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 CC Rhabditidae; Peloderinae; Caenorhabditis.  
 NC NCB1\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-Bristol N2;  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Alnscough R., Anderson K., Baynes C., Beirs M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,  
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,  
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,

RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
 RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,  
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
 RA Waterston R., Watson A., Weinstock L., Wilkison-Sproat J.,  
 RA Wohlman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP REVISIONS.  
 RA Durbin R.;  
 RL Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z11505; CAA7587.2; -;  
 DR PIR: S31130; S31130.  
 DR WormPep: F59B2.10; CE28835.  
 KW Hypothetical protein.  
 FT DOMAIN 11 14 POLY-LEU.  
 SO SEQUENCE 106 AA: 11599 MW: FFOCCF5EEC2777A6 CRC64;  
 Query Match 1.5%; Score 7; DB 1; Length 106;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 168 ISSLLIL 174  
 Db 8 ISSLLIL 14  
 RESULT 32  
 R22A\_YEAST STANDARD: PRT: 120 AA.  
 ID R22A\_YEAST  
 AC P05749;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE 60S ribosomal protein L22-A (Yl31) (Rp4).  
 GN RPL22A OR YLR061W OR L2168.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 NC NCB1\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Andre B., Urrestazu L.A.;  
 RL Submitted (JAN-1996) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 1-51.  
 RX MEDLINE=83048950; PubMed=6814480;  
 RA Otake E., Higo K.-I., Osawa S.;  
 RT "Isolation of seventeen proteins and amino-terminal amino acid  
 RT sequences of eight proteins from cytoplasmic ribosomes of yeast.";  
 RL Biochemistry 21:4545-4550(1982).  
 CC -1- MISCELLANEOUS: THERE ARE TWO GENES FOR L22 IN YEAST.  
 CC -1- SIMILARITY: BELONGS TO THE L22E FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X94607; CAA64308.1; -;

```

DR EMBL; Z73233; CAA97592.1; -.
DR PIR; S11582; S11582.
DR SGD; S0004051; RPL22A.
DR InterPro; IPR002671; Ribosomal_L22e.
DR Pfam; PF01776; Ribosomal_L22e; 1.
DR ProDom; PD007306; Ribosomal_L22e; 1.
DR Ribosomal protein; Multigene family.
KW INT_MET
FT CONFLICT 48 48 N -> D (IN REF. 2).
SQ SEQUENCE 120 AA; 13562 MW; 84050FEA84A56D59 CRC64;

Query Match
Best Local Similarity 1.5%; Score 7; DB 1; Length 120;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 DEEDEE 72
   |||||
Db 114 DEEDEE 120

RESULT 33
LSPA_CHLTR STANDARD; PRT; 167 AA.
ID LSPA_CHLTR
AC 08413;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lipoprotein signal peptidase (EC 3.4.23.36) (Prolipoprotein signal
DE peptidase) (Signal peptidase II) (Spsae II).
GN LSPA OR CT408.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UM-3/Cx;
RX MEDLINE=9900809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
CC -1- FUNCTION: THIS PROTEIN SPECIFICALLY CATALYZES THE REMOVAL OF
CC SIGNAL PEPTIDES FROM PROLIPOPROTEINS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Cleavage of N-terminal leader sequences from
CC membrane prolipoproteins. Hydrolyzes Xaa-Xbb-Xcc-|-Cys, in which
CC Xaa is hydrophobic (preferably Leu), Xbb is often Ser or Ala, Xcc
CC is often Gly or Ala, and the Cys is alkylated on sulfur with a
CC diacylglyceryl group.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001314; AAC68005.1; -.
DR MEROPS; A08.001; -.
DR InterPro; IPR001872; SigPase_A8.
DR Pfam; PF01252; Peptidase_A8; 1.
DR PRINTS; PR00781; LITOSTIGRASE.
DR ProDom; PD004304; SigPase_A8; 1.
DR TIGRfams; TIGR00077; LSPA; 1.
DR PROSITE; PS00855; SPASE_II; 1.
KW Hydrolyase; Aspartyl protease; Transmembrane; Complete proteome.
FT TRANSMEM 8 28
FT TRANSMEM 68 88 POTENTIAL.
FT TRANSMEM 101 121 POTENTIAL.

```

```

FT TRANSMEM 139 159 POTENTIAL.
FT ACT_SITE 116 116 BY SIMILARITY.
FT ACT_SITE 143 143 BY SIMILARITY.
SQ SEQUENCE 167 AA; 18939 MW; 0B9752FB80B01F12 CRC64;

Query Match
Best Local Similarity 1.5%; Score 7; DB 1; Length 167;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 420 LTLILLA 426
   |||||
Db 10 LTLILLA 16

RESULT 34
YF6_YEAST STANDARD; PRT; 169 AA.
ID YF6_YEAST
AC P43604;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical 18.8 kDa protein in HIS2-CDCl4 intergenic region.
GN YF026C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae.";
RL Nat. Genet. 10:261-268(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=96287654; PubMed=8686381;
RA Eki T., Naitou M., Hagiwara H., Abe M., Ozawa M., Sasanuma S.-I.,
RA Sasanuma M., Tsuchiya Y., Shibata T., Watanabe K., Ono A.,
RA Yamazaki M., Tashiro H., Hanaoka F., Murakami Y.;
RT "Fifteen open reading frames in a 30.8 kb region of the right arm of
RT chromosome VI from Saccharomyces cerevisiae.";
RL Yeast 12:177-190(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D50617; BAA09265.1; -.
DR SGD; S0001922; YF026C.
KW Hypothetical protein.
SQ SEQUENCE 169 AA; 18809 MW; BBlF4943C9A6815F CRC64;

Query Match
Best Local Similarity 1.5%; Score 7; DB 1; Length 169;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 ITVALLI 202
   |||||
Db 8 ITVALLI 14

RESULT 35
Y129_METUA STANDARD; PRT; 170 AA.
ID Y129_METUA

```

```

AC 057593;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein M0129.
GN M0129.
OS Methanococcus jannaschii.
OC Archaea: Euryarchaeota: Methanococci: Methanococcales:
CC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=9637999; PubMed=6688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervaege A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Nguyen D.,
RA Scott J.L., Geophagen N.S.M., Weidman J.F., Fuhrman J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Clifton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-F., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RL jannaschii."
CC Science 273:1058-1073(1996).
CC -1- SIMILARITY: STRONG. TO M.JANNASCHII M0554 AND M0587.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 067470; AAB98118.1; -
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 31 51 POTENTIAL.
FT TRANSMEM 58 78 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
SQ SEQUENCE 170 AA: 19621 MW: 4F3B62C7DDA89D9A CRC64:
QY 140 ISVIVM 146 1.5%; Score 7; DB 1; Length 170;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 144 ISVIVM 150

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Ternyn N.,
RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermair B., Meche R., Mueller M.,
RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hohnselt J., Zimmermann W., Medler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Roben J.,
RA Van der Schuren J., Grymoprez B., Chung Y.-J., Vandenbussche F.,
RA Breken M., Weltjens I., Voet M., Bastiens I., Aert R., Delfoor E.,
RA Weltzeegger T., Boche G., Ramsperger U., Hilbert H., Braun M.,
RA Holzner E., Brandt A., Peters S., van Staveren M., Dirke W.,
RA Moolman P., Klein Lankhorst R., Rose M., Hauf J., Koelter P.,
RA Bernelsier S., Hempel S., Feldpausch M., Lambers S., Van den Daele H.,
RA De Keyser A., Buysnaert C., Gleien J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,
RA Petlett A., Rajandream M.A., Lyne M., Benes V., Reckmann S.,
RA Borikova D., Bloecker H., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Grandearth K., Danner D., Herzl A.,
RA Neumann S., Argilou A., Vitale D., Liguori R., Piravandi E.,
RA Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chiford F., Cooke R., Berger C., Montfort A., Casacuberta E.,
RA Ghibons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
RA Perez-Perez A., Purrelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Belke C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Haeremann K.,
RA Patnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spielgel L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Speith J., Ryan E., Andrews S., Gelsel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
RA Ma P., Zhong J., Preston R., Vill D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hofman J., Till S.,
RA Grnath S., Shohdy N., Hasegawa A., Hamed A., Lohdi M., Johnson A.,
RA Chen E., Marra M., Martienssen R., McCombie W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana."
RN Nature 402:769-777(1999).
RP [3]
RP SEQUENCE OF 1-86 FROM N.A.
RC STRAIN=cv. Columbia;
RA Raynal M., Grellet F., Laudie M., Meyer Y., Cooke R., Delseny M.;
RL Submitted (OCT-1992) to the EMBL/Genbank/DBP databases.
CC -1- FUNCTION: MAY HAVE A STRUCTURAL ROLE TO STABILIZE THE LIPID BODY
CC DURING DESICCATION OF THE SEED BY PREVENTING COALESCENCE OF THE
CC OIL. PROBABLY INTERACTS WITH BOTH LIPID AND PHOSPHOLIPID MOETIES
CC OF LIPID BODIES. MAY ALSO PROVIDE RECOGNITION SIGNALS FOR SPECIFIC
CC LIPASE ANCHORAGE IN LIPOLYSIS DURING SEEDLING GROWTH.
CC -1- SUBCELLULAR LOCATION: SURFACE OF OIL BODIES. OLEOSINS EXIST AT A
CC MONOLAYER LIPID/WATER INTERFACE.
CC -----
CC -1- SIMILARITY: BELONGS TO THE OLEOSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X62353; CAA44225.1; -

```

```
DR EMBL: AL035523; CAB36756.1; -
DR EMBL: AL161562; CAB79423.1; -
DR EMBL: Z17738; CAA79049.1; -
DR PIR: S22143; S22143.
DR PIR: S22538; S22538.
DR InterPro: IPR00136; Oleosin.
DR Pfam: PF01277; Oleosin; 1.
DR PROSITE: PS00811; OLEOSINS; 1.
KW Oil body; Multigene family.
FT DOMAIN 1 45 POLAR.
FT REPEAT 46 117 HYDROPHOBIC.
SQ SEQUENCE 173 AA; 18569 MW; 4F718BC380105F73 CRC64;

Query Match
Best Local Similarity 1.5%; Score 7; DB 1; Length 173;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 ITVALLI 202
DB 90 ITVALLI 96

RESULT 37
OLE5_BRANA STANDARD; PRT; 175 AA.
ID OLE5_BRANA
AC P29109;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major oleosin Nap-II (Fragment).
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 43-77.
RC TISSUE=Seed;
RX MEDLINE=9111733; PubMed=1989697;
RA Murphy D.J., Keen J.N., O'Sullivan J.N., Au D.M.Y., Edwards E.-W.,
RA Jackson P.J., Cummins I., Gibbons T., Shaw C.H., Ryan A.J.;
RT "A class of amphipathic proteins associated with lipid storage bodies
RT in plants. Possible similarities with animal serum apolipoproteins.";
RL Biochim. Biophys. Acta 1088:86-94(1991).
CC -1- FUNCTION: MAY HAVE A STRUCTURAL ROLE TO STABILIZE THE LIPID BODY
CC DURING DESICCATION OF THE SEED BY PREVENTING COALESCENCE OF THE
CC OIL. PROBABLY INTERACTS WITH BOTH LIPID AND PHOSPHOLIPID MOIETIES
CC OF LIPID BODIES. MAY ALSO PROVIDE RECOGNITION SIGNALS FOR SPECIFIC
CC LIPASE ANCHORAGE IN LIPOLYSIS DURING SEEDLING GROWTH.
CC -1- SUBCELLULAR LOCATION: SURFACE OF OIL BODIES. OLEOSINS EXIST AT A
CC MONOLAYER LIPID/WATER INTERFACE.
CC -1- DEVELOPMENTAL STAGE: ACCUMULATES DURING THE DESICCATION PHASE OF
CC EMBRYO DEVELOPMENT.
CC -1- SIMILARITY: BELONGS TO THE OLEOSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X58000; CAA41064.1; -
DR InterPro: IPR00136; Oleosin.
DR Pfam: PF01277; Oleosin; 1.
DR PROSITE: PS00811; OLEOSINS; 1.
KW Seed; Oil body; Multigene family; Repeat.
FT NON_TER 1 47 POLAR.
FT DOMAIN 48 119 HYDROPHOBIC.
FT REPEAT 17 26
FT REPEAT 27 36
```

```
FT CONFLICT 49 49 V -> A (IN REF. 1; AA SEQUENCE).
FT CONFLICT 66 66 L -> A (IN REF. 1; AA SEQUENCE).
FT CONFLICT 72 72 A -> L (IN REF. 1; AA SEQUENCE).
SQ SEQUENCE 175 AA; 19349 MW; E166698285ABC8 CRC64;

Query Match
Best Local Similarity 1.5%; Score 7; DB 1; Length 173;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 ITVALLI 202
DB 92 ITVALLI 98

RESULT 38
OLE5_BRANA STANDARD; PRT; 183 AA.
ID OLE5_BRANA
AC P29109;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Oleosin Bn-V (BnV) (Fragment).
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Bienvenu;
RX MEDLINE=92379264; PubMed=1511134;
RA Keddie J., Edwards E.-W., Gibbons T., Shaw C., Murphy D.J.;
RT "Sequence of an oleosin cDNA from Brassica napus.";
RL Plant Mol. Biol. 19:1079-1083(1992).
CC -1- FUNCTION: MAY HAVE A STRUCTURAL ROLE TO STABILIZE THE LIPID BODY
CC DURING DESICCATION OF THE SEED BY PREVENTING COALESCENCE OF THE
CC OIL. PROBABLY INTERACTS WITH BOTH LIPID AND PHOSPHOLIPID MOIETIES
CC OF LIPID BODIES. MAY ALSO PROVIDE RECOGNITION SIGNALS FOR SPECIFIC
CC LIPASE ANCHORAGE IN LIPOLYSIS DURING SEEDLING GROWTH.
CC -1- SUBCELLULAR LOCATION: SURFACE OF OIL BODIES. OLEOSINS EXIST AT A
CC MONOLAYER LIPID/WATER INTERFACE.
CC -1- DEVELOPMENTAL STAGE: ACCUMULATES DURING THE DESICCATION PHASE OF
CC EMBRYO DEVELOPMENT.
CC -1- SIMILARITY: BELONGS TO THE OLEOSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X63779; CAA45313.1; -
DR PIR: S25089; S25089.
DR InterPro: IPR00136; Oleosin.
DR Pfam: PF01277; Oleosin; 1.
DR PROSITE: PS00811; OLEOSINS; 1.
KW Seed; Oil body; Multigene family; Repeat.
FT NON_TER 1 47 POLAR.
FT DOMAIN 48 119 HYDROPHOBIC.
FT REPEAT 11 20
FT REPEAT 21 30
SQ SEQUENCE 183 AA; 20286 MW; 7E791E2D4EC70A26 CRC64;

Query Match
Best Local Similarity 1.5%; Score 7; DB 1; Length 183;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 ITVALLI 202
DB 92 ITVALLI 98
```

```

RESULT 39
US10_HSVSB STANDARD; PRT: 208 AA.
ID US10_HSVSB STANDARD; PRT: 208 AA.
AC P30817;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Virion protein US10 homolog (Fragment).
OS Simian herpes B virus (Cercopithecidae herpesvirus 1) (Shbv).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
NCBI_Taxid=10325;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92113572; Pubmed=1309859;
RA Killen A.M., Harrington L., Wall L.V.M., Kelly D.C.;
RT "Nucleotide sequence analysis of a homologue of herpes simplex virus
type 1 gene US9 found in the genome of simian herpes B virus.";
RL J. Gen. Virol. 73:195-199(1992).
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER HSV-1 US10,
EHV-1 66, EHV-4 ORF3, AND VZV 64/69.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S75996; AAB21003.2; -.
DR PIR: P00278; P00278.
DR ZINC-finger.
FT ZN-FING 167 189 POTENTIAL.
FT NON_TER 1 1
SQ SEQUENCE 208 AA; 22837 MW; B8076C3AC0A25E6 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 396 ASATASG 402
|111111|
Db 202 ASATASG 208

RESULT 40
UBC_ASEFM2 STANDARD; PRT: 213 AA.
ID UBC_ASEFM2 STANDARD; PRT: 213 AA.
AC P25869;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-21 kDa (EC 6.3.2.19)
DE (ubiquitin-protein ligase) (ubiquitin carrier protein).
GN ASFV-UBC.
OS African swine fever virus (isolate Malawi Lil 20/1) (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
NCBI_Taxid=10500;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92155177; Pubmed=1310934;
RA Hingamp P.M., Arnold J.E., Mayer R.J., Dixon L.K.;
RT "A ubiquitin conjugating enzyme encoded by African swine fever
virus.";
RL EMBO J. 11:361-366(1992).
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
OTHER PROTEINS.
CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +
diphosphate + protein N-ubiquityllysine.
CC -1- PATHWAY: Ubiquitin conjugation; second step.

```

```

CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC UBIQUITIN-THIOLESTER FORMATION.
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X62440; CAA44305.1; -.
DR EMBL: X71982; CAA50851.1; -.
DR PIR: S19158; S19158.
DR HSPSP; 002159; 20CZ.
DR InterPro: IPR000608; UBQ_conjugat.
DR Pfam: PF00179; UQ_con; 1.
DR ProDom: PD000461; UBQ_conjugat; 1.
DR SMART: SM00212; UBQC; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS50127; UBIQUITIN_CONJUGAT_2; 1.
DR SMART: SM00212; UBQC; 1.
KW Ub1 conjugation pathway; Ligase.
FT BINDING 85 85 UBIQUITIN (BY SIMILARITY).
FT DOMAIN 183 213 ASP/GLU-RICH (ACIDIC).
SQ SEQUENCE 213 AA; 24468 MW; F9671BC7385D6DE CRC64;

Query Match 1.5%; Score 7; DB 1; Length 213;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 66 DEEDDEE 72
|111111|
Db 203 DEEDDEE 209

RESULT 41
RADC_SYNY3 STANDARD; PRT: 243 AA.
ID RADC_SYNY3 STANDARD; PRT: 243 AA.
AC P52601;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA repair protein radc homolog.
GN RADC OR SL0766.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
NCBI_Taxid=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; Pubmed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
CC -1- FUNCTION: INVOLVED IN DNA REPAIR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE RADC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D64000; BAA10123.1; -.
DR InterPro: IPR003583; HHH_1.
DR InterPro: IPR001405; RadC.
DR ProDom: PD007415; RadC; 1.
DR SMART: SM00278; Hhhl; 1.

```



DR TIGRFAMS: TIGR00608; radc: 1.  
 DR PROSITE: PS01302; RADC: 1.  
 DR DNA repair: Complete proteome.  
 SO SEQUENCE 243 AA; 26973 MW; F0E9D51F999EF6B4 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 243;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 75 LKYGAKH 81  
 |||||  
 Db 21 LKYGAKH 27

## RESULT 42

ISPE.THEMA STANDARD; PRT: 271 AA.  
 ID ISPE.THEMA  
 AC 09XIA3;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (EC 2.7.1.148) (CMK)  
 DE (4-(cytidine-5'-diphospho)-2-C-methyl-D-erythritol kinase).  
 GN ISPE OR TM1383.  
 OS Thermotoga maritima.  
 OC Bacteria; Thermotogae; Thermotogales;  
 OC Thermotogaceae; Thermotoga.  
 OX NCBI\_TaxID=2336;

RP SEQUENCE FROM N.A.  
 RC STRAIN=MSB / DSM 3109;  
 RX MEDLINE=99287316; PubMed=10360571;  
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,  
 Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,  
 McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,  
 Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,  
 Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,  
 Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;  
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from  
 genome sequence of Thermotoga maritima."  
 RL Nature 399:323-329(1999).

RT Nature 399:323-329(1999).  
 RL Nature 399:323-329(1999).  
 CC -1- FUNCTION: CATALYZES THE PHOSPHORYLATION OF THE POSITION 2 HYDROXY  
 GROUP OF 4-DIPHOSPHOCYTIDYL-2C-METHYL-D-ERYTHRITOL (BY  
 SIMILARITY).

CC -1- CATALYTIC ACTIVITY: ATP + 4-(cytidine 5'-diphospho)-2-C-methyl-D-  
 erythritol = ADP + 2-phospho-4-(cytidine 5'-diphospho)-2-C-methyl-D-  
 erythritol.  
 CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; fourth  
 step.

CC -1- SIMILARITY: BELONGS TO THE ISPE FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AE001791; AAD36453.1; -.  
 DR TIGR: TM1383; -.  
 DR InterPro: IPR001745; GHMPKase\_ATP.  
 DR InterPro: IPR004424; ISPE.  
 DR Pfam: PF00288; GHMP\_Kinases; 1.  
 DR TIGRFAMS: TIGR00154; ispe.1.  
 KW Transferrase; Kinase; Isoprene biosynthesis; ATP-binding;  
 KW Complete proteome.  
 FT NP\_BIND 97 107 ATP (POTENTIAL).  
 SO SEQUENCE 271 AA; 30056 MW; 3E9E47FDAB0F336D CRC64;

Query Match 1.5%; Score 7; DB 1; Length 271;  
 Best Local Similarity 100.0%; Pred. No. 58;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 181 YLGEVERK 187  
 |||||  
 Db 115 YLGEVERK 121

## RESULT 43

YITL\_BACSU STANDARD; PRT: 280 AA.  
 ID YITL\_BACSU  
 AC P39803; P70946; 008139;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein yitl.  
 GN YITL.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;

RP SEQUENCE FROM N.A.  
 RC STRAIN=168 / DB104;  
 RX MEDLINE=94042884; PubMed=8226659;  
 RA Shiga Y., Yamagata H., Ueda S.;  
 RT "Characterization of the gene encoding an intracellular proteinase  
 RT inhibitor of Bacillus subtilis and its role in regulation of the  
 RT major intracellular proteinase."  
 RL J. Bacteriol. 175:7130-7137(1993).

RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=97177785; PubMed=9025291;  
 RA Levine A., Vannier F., Roche B., Autret S., Mavel D., Seror S.J.;  
 RT "A 10.3 kbp segment from nprB to argJ at the 102 degrees region of the  
 RT Bacillus subtilis chromosome."  
 RL Microbiology 143:175-177(1997).

RP SEQUENCE FROM N.A.  
 RX MEDLINE=98015415; PubMed=9353931;  
 RA Medina N., Vannier F., Roche B., Autret S., Levine A., Seror S.J.;  
 RT "Sequencing of regions downstream of addA (98 degrees) and cItg (289  
 RT degrees) in Bacillus subtilis."  
 RL Microbiology 143:3305-3308(1997).

RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 Azevedo V., Bertolo R., Bessieres P., Bolotin A., Borchert S.,  
 Borriis R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,  
 Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
 Enlart K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 Fritz C., Fujita M., Fujita Y., Funo S., Gallzer A., Galleron N.,  
 Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
 Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
 Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
 Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,  
 Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,  
 Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,  
 Parco V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,  
 Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 Sekiguchi Y., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
 Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 Tanouchi M., Tamakoshi A., Tanaka T., Terpsita P., Tognoni A.,  
 Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vasseroiti A.,  
 Viari A., Wambutt R., Wedler E., Wedler H., Weltzenegger T.,  
 Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,



```

RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: TO B.SUBTILIS YPOC AND YOFU.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D16311; BAA03818.1; -
DR EMBL: Y08476; CAI70630.1; -
DR EMBL: Z79580; CAB01834.1; -
DR EMBL: Z99109; CAB12952.1; -
DR PIR: A49921; A49921.
DR Subtilisin: BG10850; YLT.
DR InterPro: IPR003740; DUF161.
DR Pfam: PF02588; DUF161.1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 9 29 POTENTIAL.
FT TRANSMEM 54 74 POTENTIAL.
FT TRANSMEM 80 100 POTENTIAL.
FT TRANSMEM 151 171 POTENTIAL.
FT CONFLICT 121 121 V -> D (IN REF. 1).
FT CONFLICT 226 226 V -> L (IN REF. 1).
FT CONFLICT 278 278 R -> T (IN REF. 1).
SQ SEQUENCE 280 AA; 30535 MW; 4BF98B12B8DA59E7 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 280;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 73 LTLKYGA 79
Db 125 LTLKYGA 131

RESULT 44
MCM1_YEAST
ID MCM1_YEAST STANDARD; PRT; 286 AA.
AC P11746.
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Pheromone receptor transcription factor (GRM/PRTF protein).
DE MCM1 OR FUN80 OR YMR043W OR YMR532.08.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
ON [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90249735; Pubmed=2159934;
RA Ammerer G.;
RT "Identification, purification, and cloning of a polypeptide
RT (PRTF/GRM) that binds to mating-specific promoter elements in
RT yeast."
RL Genes Dev. 4:299-312(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89141759; Pubmed=3066908;
RA Passmore S., Maine G.T., Elble R., Christ C., Tye B.K.;
RT "Saccharomyces cerevisiae protein involved in plasmid maintenance is
RT necessary for mating of MAT alpha cells."
RL J. Mol. Biol. 204:593-606(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88030692; Pubmed=3311883;

```

```

RA Dubois E., Bercy J., Descamps F., Messenguy F.;
RT "Characterization of two new genes essential for vegetative growth in
RT Saccharomyces cerevisiae: nucleotide sequence determination and
RT chromosome mapping."
RL Gene 55:265-275(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=8288c / AB972;
RA Odell C., Bowman S., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-1995) to the EMBL/Genbank/DBJ databases.
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF COMPLEX WITH MAT2.
RX MEDLINE=98140702; Pubmed=9490409;
RA Tan S., Richmond T.J.;
RT "Crystal structure of the yeast Matalpha2/Mcm1/DNA ternary complex."
RL Nature 391:660-666(1998).
CC -1- FUNCTION: INTERACTS WITH THE ALPHA-2 REPRESSOR OR WITH THE
CC ALPHA-1 ACTIVATOR THEREBY REGULATING THE EXPRESSION OF MATING-
CC TYPE-SPECIFIC GENES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X52453; CAA36691.1; -
DR EMBL: X14187; CAA32389.1; -
DR EMBL: M17511; AAA34609.1; -
DR EMBL: Z48502; CAA88409.1; -
DR PIR: A34599; A34599.
DR PDB: 1MNM; 18-MAR-98.
DR TRANSFAC: T00500; -
DR TRANSFAC: T00501; -
DR SGD: S0004646; MCM1.
DR InterPro: IPR002100; TF_MADSbox.
DR Pfam: PF00319; SRF-TF.1.
DR PRINTS: PR00404; MADSDOMAIN.
DR SMART: SM00432; MADS.1.
DR PROSITE: PS00350; MADS_BOX_1; 1.
DR PROSITE: PS50066; MADS_BOX_2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein; 3D-structure.
FT DOMAIN 18 72 MADS.
FT CONFLICT 9 9 ASP/GLU-RICH (ACIDIC).
FT CONFLICT 37 37 P -> S (IN REF. 3).
FT CONFLICT 136 137 GA -> F (IN REF. 3).
FT CONFLICT 158 286 GA -> AR (IN REF. 3).
SQ SEQUENCE 286 AA; 32802 MW; FD75B4BF549E9E3B CRC64;

Query Match 1.5%; Score 7; DB 1; Length 286;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 66 DEEDDEE 72
Db 99 DEEDDEE 105

RESULT 45
ZIPA_VIBCH
ID ZIPA_VIBCH STANDARD; PRT; 291 AA.
AC Q9KTD2.
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cell division protein zipa homolog.
RN [1]
RP ZIPA OR VC0970.

```

```

OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tetelin H., Richardson D.,
RA Emdonlaeva M.D., Vamathevan J., Bass S., Olin H., Dragoi I., Sellers P.,
RA Salzman L., Ueberlack T., Fleischmann R.D., Nieman W.C., White O.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000)
CC -!- FUNCTION: Interacts directly with the cell division protein ftsZ.
CC Probable receptor for the septal ring structure, may anchor it
CC to the inner-membrane (By similarity).
CC -!- SUBCELLULAR LOCATION: Type IB membrane protein. Inner membrane (By
CC similarity).
CC -!- SIMILARITY: BELONGS TO THE ZIPA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE004178; AAF94132.1; -.
DR HSSP: P71713; 1F7X.
DR TIGR: VC0970; -.
KW Cell division; Septation; Transmembrane; Inner membrane;
KW Complete proteome.
FT DOMAIN 1 5 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 6 26 POTENTIAL.
FT DOMAIN 27 291 CYTOPLASMIC (POTENTIAL).
SO SEQUENCE 291 AA; 32786 MW; A7DE34AC3DE64A35 CRC64;

Query Match 1.58; Score 7; DB 1; Length 291;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 DEEDDE 72
1 |
DB 94 DEEDDE 100

RESULT 46
SAPC.ECOLI STANDARD; PRT; 296 AA.
AC 047624;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peptide transport system permease protein sapc.
GN SAPC OR B1292.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / FRAG5;
RA Epstein W., Noelker E., Stumpe S., Tewes R., Schmid R., Bakker E.P.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;

```

```

RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA Sempel G., Seki Y., Sivasubraman S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map."
RL DNA Res. 3:363-377(1996).
CC -!- FUNCTION: INVOLVED IN A PEPTIDE INTAKE TRANSPORT SYSTEM THAT
CC PLAYS A ROLE IN THE RESISTANCE TO ANTIMICROBIAL PEPTIDES.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -!- SIMILARITY: BELONGS TO THE BINDING-PROTEIN-DEPENDENT TRANSPORT
CC SYSTEM PERMEASE FAMILY. OPBP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X97282; CAA65939.1; -.
DR EMBL: AE000227; AAC74374.1; -.
DR EMBL: D90767; BAA14853.1; -.
DR EMBL: D90768; BAA14862.1; -.
DR EMBL: D90766; BAA14845.1; -.
DR EcoGene: EG20256; sapc.
DR InterPro: IPR000515; BPD_transp.
DR Pfam: PF00528; BPD_transp. 1.
DR PROSITE: PS00402; BPD_TRANS_P. INN_MEMBR. FALSE_NEG.
KW Transp. Peptide transport; Transmembrane; Inner membrane;
KW Complete proteome.
FT TRANSMEM 30 50 POTENTIAL.
FT TRANSMEM 99 119 POTENTIAL.
FT TRANSMEM 134 154 POTENTIAL.
FT TRANSMEM 197 217 POTENTIAL.
FT TRANSMEM 223 243 POTENTIAL.
FT TRANSMEM 259 279 POTENTIAL.
SO SEQUENCE 296 AA; 31548 MW; 57CD1F8180BF0B7C CRC64;

Query Match 1.58; Score 7; DB 1; Length 296;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 AATMISV 142
DB 271 AATMISV 277

RESULT 47
NKX1.BISBI STANDARD; PRT; 300 AA.
AC 046383;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sodium/potassium/calcium exchanger 1 (Na(+)/K(+/Ca(2+)-
DE exchange protein 1) (Retinal rod Na-Ca+K exchanger) (Fragment).
GN SLC24A1 OR NKX1.
OS Bison (American bison).

```

```

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
CC Ovis; Bovinae; Bison.
CC NCBI_TaxID=9901;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=98138491; PubMed=9478004;
RA Tucker J.E., Winkfein R.J., Cooper C.B., Schneckamp P.P.;
RT "cDNA cloning of the human retinal rod Na-Ca + K exchanger: comparison
RT with a revised bovine sequence."
RL Invest. Ophthalmol. Vis. Sci. 39:435-440(1998).
CC -1- FUNCTION: Critical component of the visual transduction cascade,
CC controlling the calcium concentration of outer segments during
CC light and darkness. Light causes a rapid lowering of cytosolic
CC free calcium in the outer segment of both retinal rod and cone
CC photoreceptors and the light-induced lowering of calcium is caused
CC by extrusion via this protein which plays a key role in the
CC process of light adaptation. Transports one Ca(2+) and one K(+) in
CC exchange for four Na(+) (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE SLC24A FAMILY OF TRANSPORTERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF025480; AAC13320.1; -.
KW Vision; Transport; Antiport; Symport; Calcium transport;
KW Potassium transport; Sodium transport; Transmembrane.
FT NON_TER 1 1
FT TRANSMEM 259 275 POTENTIAL.
FT DOMAIN 218 242 POLY-GLU.
FT NON_TER 300 300
SQ SEQUENCE 300 AA; 31671 MW; 2BE592DA5A89781E CRC64;

Query Match 1.5%; Score 7; DB 1; Length 300;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 66 DEEDDEE 72
Db 224 DEEDDEE 230

RESULT 48
D7_AEDAE STANDARD; PRT; 321 AA.
AC P18153;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE D7 protein precursor (Allergen Aed a 2).
GN D7
OS Aedes aegypti (Yellowfever mosquito).
CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
CC Culicoidae; Aedes.
OX NCBI_TaxID=7159;
RN 11
RP SEQUENCE FROM N.A.
RX STRAIN=Rockefeller; TISSUE=Salivary gland;
RX MEDLINE=91270277; PubMed=2052024;
RA James A.A., Blackmer K., Marinotti O., Ghosh C.R., Racioppi J.V.;
RT "Isolation and characterization of the gene expressing the major
RT salivary gland protein of the female mosquito, Aedes aegypti."
RL Mol. Biochem. Parasitol. 44:245-254(1991).
CC -1- FUNCTION: THOUGHT TO BE INVOLVED IN BLOOD-FEEDING.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: ADULT FEMALE SALIVARY GLAND.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M3156; AAA29347.1; -.
DR EMBL: M3157; AAA29348.1; -.
KW Allergen; Signal.
FT SIGNAL 1 17
FT CHAIN 18 321
FT VARIANT 4 4
FT VARIANT 10 10
FT VARIANT 19 19
FT VARIANT 30 30
FT VARIANT 43 43
FT VARIANT 115 115
FT VARIANT 159 159
FT VARIANT 169 169
FT VARIANT 192 192
FT VARIANT 221 221
FT VARIANT 223 223
FT VARIANT 226 226
FT VARIANT 234 235
FT VARIANT 259 259
FT VARIANT 301 301
SQ SEQUENCE 321 AA; 37028 MW; AA8D6328E46FBC9 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 321;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 422 LLLAIF 428
Db 4 LLLAIF 10

RESULT 49
RAD1_SCHPO STANDARD; PRT; 323 AA.
AC P22193;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA repair protein rad1.
GN RAD1 OR SPAC1952.07.
OS Schizosaccharomyces pombe (fission yeast).
CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
CC Schizosaccharomycetales; Schizosaccharomycetaceae;
CC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN 11
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RX MEDLINE=95011649; PubMed=7926829;
RA Long K.E., Sunnerhagen P., Subramani S.;
RT "The Schizosaccharomyces pombe rad1 gene consists of three exons and
RT the cDNA sequence is partially homologous to the Ustilago maydis REC1
RT cDNA."
RL Gene 148:155-159(1994).
RN 12
RP SEQUENCE FROM N.A.
RX STRAIN=90287166; PubMed=2355921;
RX Sunnerhagen P., Seaton B.L., Nasim A., Subramani S.;
RT "Cloning and analysis of a gene involved in DNA repair and
RT recombination, the rad1 gene of Schizosaccharomyces pombe."
RL Mol. Cell. Biol. 10:3750-3760(1990).
RN 13
RP REVISIONS.
RX MEDLINE=94261448; PubMed=8202381;

```

RA Zhang M.O., Marr T.G.;  
RT "Fission yeast gene structure and recognition."  
RL Nucleic Acids Res. 22:1750-1759(1994).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-972;  
RX MEDLINE-21848401; Pubmed-11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Squires J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,  
RA Brooke K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA Jones K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Shelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volckaert G., Aert R., Robben J., Gymnopoulos B.,  
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gebel C., Fuchs M., Filiz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
RA Beyer P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Gilbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rocher M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shipkova I., G.V., Ussery D., Barrett B.G., Nurse P.;  
RT The genome sequence of Schizosaccharomyces pombe.  
RL Nature 415:871-880(2002).  
CC -1- FUNCTION: RECOMBINATIONAL REPAIRMENT OF DNA DAMAGES INDUCED BY UV  
AND GAMMA RADIATION. COULD HAVE EXONUCLEASE ACTIVITY.  
CC -1- ENZYME REGULATION: INHIBITED BY CAFFEINE.  
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).  
CC -1- SIMILARITY: SOME, TO HUMAN AND HAMSTER GADD45.  
CC -1- SIMILARITY: SOME, YEAST RAD17 AND TO U. MAYDIS RECL.  
CC -1- CAUTION: REF. 2 SEQUENCE STARTED AT POSITION 130, DUE TO THE FACT  
THAT TWO EXONS WERE MISSED.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; 228385; CAAG2231.1; -  
CC EMBL; M38132; AAA35329.1; -  
CC EMBL; ALI09820; CAB52571.1; -  
CC PIR; A35680; A35680.  
DR InterPro: IPR003021; Rad1\_Recl.  
DR Pfam: PF02144; Rad1.1.  
DR PRINTS; PR01245; RADIRECL.  
KW DNA damage; DNA repair; Hydrolyase; Exonuclease; Nuclear protein.  
FT DOMAIN 285 294 ASP/GLU-RICH (ACIDIC).  
FT FT 285 323 ASP/GLU-RICH (ACIDIC).  
SQ SEQUENCE 323 AA; 36527 MW; 1A2ADDFC3A55EB6 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 323;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 66 DEEDED 72  
Db 286 DEEDED 292

RESULT 50  
1079\_STYNY3

ID YE09\_STYNY3 STANDARD; PRT; 326 AA.  
AC P73594;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical WD-repeat protein slr1409.  
GN SLR1409.  
OS Synechocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
OX NCBI\_TaxID=1148;  
RA [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-97061201; Pubmed-8905231;  
RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hikosawa M., Sugita M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naito K., Okumura S.,  
RA Shimp S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
RA Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions."  
RL DNA Res. 3:109-136(1996).  
CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; D90907; BAA17638.1; -  
CC InterPro: IPR001680; WD40.  
CC Pfam: PF00400; WD40; 7.  
DR PRINTS; PR00320; GPROTEINRPT.  
DR PRODOM; PD000018; WD40; 1.  
DR SMART; SM00320; WD40; 7.  
DR PROSITE; PS00678; WD\_REPEATS\_1; 2.  
DR PROSITE; PS50082; WD\_REPEATS\_2; 4.  
DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
KW Hypothetical protein; Repeat; WD repeat; Complete proteome.  
FT REPEAT 47 77 WD 1.  
FT REPEAT 88 118 WD 2.  
FT REPEAT 129 159 WD 3.  
FT REPEAT 169 199 WD 4.  
FT REPEAT 210 240 WD 5.  
FT REPEAT 252 282 WD 6.  
SQ SEQUENCE 326 AA; 35759 MW; BD83A2403DE163D CRC64;

Query Match 1.5%; Score 7; DB 1; Length 326;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 109 KDQGLIV 115  
Db 283 KDQGLIV 289

Search completed: February 6, 2003, 14:19:28  
Job time: 47.4398 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 14:21:16 : Search time 17.4398 Seconds  
(Without alignments)  
593.645 Million cell updates/sec

Title: US-09-689-159a-2

Perfect score: 467  
Sequence: 1 MFELPAPLSTVFQNAQMSDN.....ATDYLVPFMDLAFHQFYI 467

Scoring table: OLIGO  
Gapop 60.0 , Gapect 60.0

Searched: 129505 seqs, 22169297 residues

Word size : 0

Total number of hits satisfying chosen parameters: 129505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

Published Applications-AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	366	78.4	467	10 US-09-754-949-4	Sequence 4, Appl1
2	366	78.4	467	10 US-09-878-454A-3	Sequence 3, Appl1
3	366	78.4	467	10 US-09-895-035-14	Sequence 14, Appl1
4	366	78.4	467	12 US-10-071-900-1	Sequence 1, Appl1
5	337	72.2	467	10 US-09-895-035-12	Sequence 12, Appl1
6	265	56.7	467	10 US-09-875-474-2	Sequence 2, Appl1
7	204	43.7	467	10 US-09-785-474-4	Sequence 4, Appl1
8	204	43.7	467	10 US-09-785-474-28	Sequence 28, Appl1
9	204	43.7	467	10 US-09-785-474-30	Sequence 30, Appl1
10	204	43.7	467	10 US-09-785-474-32	Sequence 32, Appl1
11	132	28.3	180	10 US-09-895-035-1	Sequence 1, Appl1
12	128	27.4	354	10 US-09-823-153-8	Sequence 8, Appl1
13	48	10.3	65	10 US-09-864-761-38056	Sequence 38056, A
14	44	9.4	101	10 US-09-925-299-1227	Sequence 1227, Ap
15	28	6.0	131	9 US-10-051-767-10	Sequence 10, Appl1
16	28	6.0	448	10 US-09-754-949-6	Sequence 6, Appl1
17	28	6.0	448	10 US-09-878-454A-1	Sequence 1, Appl1
18	28	6.0	448	12 US-10-071-900-2	Sequence 2, Appl1
19	22	4.7	74	9 US-10-051-767-9	Sequence 9, Appl1

20	4.7	74	9	US-10-051-767-15	Sequence 15, Appl1
21	4.7	129	9	US-10-051-767-11	Sequence 11, Appl1
22	3.9	166	9	US-10-051-767-7	Sequence 7, Appl1
23	3.9	166	9	US-10-051-767-13	Sequence 13, Appl1
24	3.9	210	10	US-10-051-767-12	Sequence 12, Appl1
25	3.4	17	10	US-09-823-153-6	Sequence 6, Appl1
26	3.2	15	10	US-09-823-153-9	Sequence 9, Appl1
27	3.2	16	10	US-09-823-153-5	Sequence 5, Appl1
28	1.7	8	10	US-08-795-903A-9	Sequence 9, Appl1
29	1.7	8	10	US-09-796-264-9	Sequence 9, Appl1
30	1.7	8	10	US-09-845-226-9	Sequence 9, Appl1
31	1.7	45	10	US-09-982-172-103	Sequence 103, App
32	1.7	45	10	US-09-982-172-227	Sequence 227, App
33	1.7	263	9	US-10-004-717-7	Sequence 7, Appl1
34	1.7	263	9	US-10-004-717-31	Sequence 31, Appl1
35	1.7	263	9	US-10-004-717-38	Sequence 38, Appl1
36	1.7	1272	10	US-09-769-097-2	Sequence 2, Appl1
37	1.7	1272	10	US-09-769-097-4	Sequence 4, Appl1
38	1.7	1276	10	US-09-866-866A-8	Sequence 8, Appl1
39	1.7	1280	9	US-10-072-621-7	Sequence 7, Appl1
40	1.7	1280	9	US-10-044-671-2	Sequence 2, Appl1
41	1.7	1280	10	US-09-866-866A-2	Sequence 2, Appl1
42	1.7	1280	10	US-09-866-866A-4	Sequence 4, Appl1
43	1.5	8	10	US-09-795-903A-6	Sequence 6, Appl1
44	1.5	8	10	US-09-796-264-6	Sequence 6, Appl1
45	1.5	8	10	US-09-845-226-6	Sequence 6, Appl1
46	1.5	25	10	US-09-864-761-38169	Sequence 38169, A
47	1.5	55	10	US-09-864-761-38937	Sequence 38937, A
48	1.5	58	10	US-09-764-869-1215	Sequence 1215, Ap
49	1.5	64	10	US-09-814-122-39	Sequence 39, Appl1
50	1.5	67	10	US-09-864-761-47359	Sequence 47359, A
51	1.5	70	10	US-09-864-761-34902	Sequence 34902, A
52	1.5	74	10	US-09-071-838-56	Sequence 56, Appl1
53	1.5	118	10	US-09-897-898-8	Sequence 8, Appl1
54	1.5	118	10	US-09-897-898-16	Sequence 16, Appl1
55	1.5	119	9	US-09-992-593A-165	Sequence 165, App
56	1.5	119	9	US-09-989-293A-165	Sequence 165, App
57	1.5	119	9	US-10-063-547-76	Sequence 26, Appl1
58	1.5	119	9	US-09-989-734-165	Sequence 165, App
59	1.5	119	9	US-09-990-444-165	Sequence 165, App
60	1.5	119	9	US-09-989-730-165	Sequence 165, App
61	1.5	119	9	US-09-990-436-165	Sequence 165, App
62	1.5	119	9	US-09-991-181-165	Sequence 165, App
63	1.5	119	9	US-09-993-687-165	Sequence 165, App
64	1.5	119	9	US-09-989-734-165	Sequence 165, App
65	1.5	119	9	US-09-997-653-165	Sequence 165, App
66	1.5	119	9	US-10-174-590-140	Sequence 140, App
67	1.5	119	9	US-10-176-758-140	Sequence 140, App
68	1.5	119	9	US-10-063-616-26	Sequence 26, Appl1
69	1.5	119	9	US-10-175-737-140	Sequence 140, App
70	1.5	119	9	US-09-993-687-165	Sequence 165, App
71	1.5	119	9	US-10-053-502-26	Sequence 26, Appl1
72	1.5	119	9	US-10-173-706-140	Sequence 140, App
73	1.5	119	9	US-10-175-752-140	Sequence 140, App
74	1.5	119	9	US-10-176-482-140	Sequence 140, App
75	1.5	119	9	US-10-176-482-140	Sequence 140, App
76	1.5	119	9	US-10-176-757-140	Sequence 140, App
77	1.5	119	9	US-10-176-913-140	Sequence 140, App
78	1.5	119	9	US-10-180-552-140	Sequence 140, App
79	1.5	119	10	US-09-989-722-165	Sequence 165, App
80	1.5	119	10	US-09-989-723-165	Sequence 165, App
81	1.5	119	10	US-09-989-723-165	Sequence 165, App
82	1.5	119	10	US-09-989-727-165	Sequence 165, App
83	1.5	119	10	US-09-989-731-165	Sequence 165, App
84	1.5	119	10	US-09-989-732-165	Sequence 165, App
85	1.5	119	10	US-09-991-073-165	Sequence 165, App
86	1.5	119	10	US-09-990-442-165	Sequence 165, App
87	1.5	119	10	US-09-991-163-165	Sequence 165, App
88	1.5	119	10	US-09-993-604-165	Sequence 165, App
89	1.5	119	10	US-09-990-456-165	Sequence 165, App
90	1.5	119	10	US-09-989-721-165	Sequence 165, App
91	1.5	119	10	US-10-006-867-26	Sequence 26, Appl1
92	1.5	119	12		

93	7	1.5	119	12	US-10-052-598-140	Sequence 140, App	166	7	1.5	774	10	US-09-745-763-115	Sequence 115, App
94	7	1.5	232	10	US-09-897-586-11	Sequence 11, Appl	167	7	1.5	896	9	US-09-903-170C-5	Sequence 5, Appl1
95	7	1.5	242	10	US-09-864-761-36180	Sequence 36180, A	168	7	1.5	896	10	US-09-903-180B-5	Sequence 5, Appl1
96	7	1.5	247	9	US-09-738-626-6438	Sequence 6438, Ap	169	7	1.5	896	10	US-09-903-171A-5	Sequence 5, Appl1
97	7	1.5	264	9	US-09-510-332-47	Sequence 47, Appl	170	7	1.5	896	10	US-09-903-188A-5	Sequence 5, Appl1
98	7	1.5	286	10	US-09-801-368-188	Sequence 188, App	171	7	1.5	896	10	US-09-903-323A-5	Sequence 5, Appl1
99	7	1.5	301	9	US-09-738-626-5030	Sequence 5030, Ap	172	7	1.5	896	10	US-09-903-325A-5	Sequence 5, Appl1
100	7	1.5	312	9	US-09-510-332-39	Sequence 39, Appl	173	7	1.5	906	9	US-10-060-230-19	Sequence 19, Appl1
101	7	1.5	319	9	US-09-738-626-5702	Sequence 5702, Ap	174	7	1.5	906	9	US-10-060-230-20	Sequence 20, Appl
102	7	1.5	341	9	US-09-738-626-4828	Sequence 4828, Ap	175	7	1.5	906	9	US-10-060-230-21	Sequence 21, Appl
103	7	1.5	366	10	US-09-887-569A-2	Sequence 2, Appl1	176	7	1.5	906	9	US-10-060-230-22	Sequence 22, Appl
104	7	1.5	392	10	US-09-902-684-2	Sequence 2, Appl1	177	7	1.5	979	10	US-09-903-187A-5	Sequence 5, Appl1
105	7	1.5	405	10	US-09-902-684-13	Sequence 13, Appl	178	7	1.5	2448	10	US-09-815-242-11870	Sequence 11870, A
106	7	1.5	415	10	US-09-712-363-237	Sequence 237, App	179	7	1.3	17	9	US-09-981-876-221	Sequence 221, App
107	7	1.5	451	10	US-09-897-898-19	Sequence 19, Appl	180	6	1.3	17	10	US-09-864-761-38208	Sequence 38208, A
108	7	1.5	494	9	US-09-902-525-52	Sequence 52, Appl	181	6	1.3	18	10	US-09-864-761-35295	Sequence 35295, A
109	7	1.5	527	9	US-09-962-678-2	Sequence 2, Appl1	182	6	1.3	19	10	US-09-864-761-43691	Sequence 43691, A
110	7	1.5	527	9	US-09-981-353-166	Sequence 166, App	183	6	1.3	19	10	US-09-734-520-16	Sequence 16, Appl
111	7	1.5	527	9	US-09-854-133-216	Sequence 216, App	184	6	1.3	19	10	US-09-071-838-232	Sequence 232, App
112	7	1.5	527	9	US-10-174-590-522	Sequence 522, App	185	6	1.3	19	12	US-10-012-034A-16	Sequence 16, Appl
113	7	1.5	527	9	US-10-176-758-522	Sequence 522, App	186	6	1.3	20	9	US-09-736-457-1855	Sequence 1855, Ap
114	7	1.5	527	9	US-10-175-737-522	Sequence 522, App	187	6	1.3	20	9	US-09-736-457-1855	Sequence 1855, Ap
115	7	1.5	527	9	US-10-173-706-522	Sequence 522, App	188	6	1.3	20	9	US-09-902-941-1855	Sequence 1855, Ap
116	7	1.5	527	9	US-10-175-738-522	Sequence 522, App	189	6	1.3	20	9	US-09-902-941-1855	Sequence 1855, Ap
117	7	1.5	527	9	US-10-175-752-522	Sequence 522, App	190	6	1.3	20	9	US-09-902-941-1855	Sequence 1855, Ap
118	7	1.5	527	9	US-10-176-482-522	Sequence 522, App	191	6	1.3	20	9	US-09-849-626-1855	Sequence 1855, Ap
119	7	1.5	527	9	US-10-176-757-522	Sequence 522, App	192	6	1.3	22	10	US-09-849-626-1856	Sequence 1856, Ap
120	7	1.5	527	9	US-10-176-913-522	Sequence 522, App	193	6	1.3	25	9	US-09-912-672A-7	Sequence 28, Appl
121	7	1.5	527	9	US-10-180-552-522	Sequence 522, App	194	6	1.3	25	10	US-09-911-888B-35	Sequence 35, Appl
122	7	1.5	527	9	US-10-180-557-522	Sequence 522, App	195	6	1.3	27	9	US-09-983-802-642	Sequence 642, App
123	7	1.5	527	10	US-09-738-973-216	Sequence 216, App	196	6	1.3	29	10	US-09-917-340-31	Sequence 31, Appl
124	7	1.5	527	12	US-10-052-586-522	Sequence 522, App	197	6	1.3	33	9	US-09-842-582-6	Sequence 6, Appl1
125	7	1.5	566	9	US-09-992-598-41	Sequence 41, Appl	198	6	1.3	39	10	US-09-925-299-5375	Sequence 1515, Ap
126	7	1.5	566	9	US-09-989-293A-41	Sequence 41, Appl	199	6	1.3	40	9	US-09-738-626-5512	Sequence 5372, Ap
127	7	1.5	566	9	US-09-989-735-41	Sequence 41, Appl	200	6	1.3	41	10	US-09-764-878-148	Sequence 148, App
128	7	1.5	566	9	US-09-990-444-41	Sequence 41, Appl	201	6	1.3	42	10	US-09-864-761-46270	Sequence 46270, A
129	7	1.5	566	9	US-09-989-730-41	Sequence 41, Appl	202	6	1.3	42	10	US-09-764-847-656	Sequence 656, App
130	7	1.5	566	9	US-09-990-436-41	Sequence 41, Appl	203	6	1.3	44	10	US-09-864-761-41268	Sequence 41268, A
131	7	1.5	566	9	US-09-991-181-41	Sequence 41, Appl	204	6	1.3	46	9	US-09-864-761-38584	Sequence 38584, A
132	7	1.5	566	9	US-09-993-687-41	Sequence 41, Appl	205	6	1.3	48	9	US-09-983-802-278	Sequence 278, App
133	7	1.5	566	9	US-09-989-734-41	Sequence 41, Appl	206	6	1.3	48	9	US-09-981-876-234	Sequence 234, App
134	7	1.5	566	9	US-09-997-653-41	Sequence 41, Appl	207	6	1.3	50	10	US-09-864-761-33963	Sequence 33963, A
135	7	1.5	566	9	US-10-174-590-56	Sequence 56, Appl	208	6	1.3	53	9	US-09-981-876-157	Sequence 157, App
136	7	1.5	566	9	US-10-176-758-56	Sequence 56, Appl	209	6	1.3	53	10	US-09-764-877-1118	Sequence 1118, Ap
137	7	1.5	566	9	US-10-175-737-56	Sequence 56, Appl	210	6	1.3	54	10	US-09-864-761-36621	Sequence 36621, A
138	7	1.5	566	9	US-09-993-667-41	Sequence 41, Appl	211	6	1.3	55	10	US-09-774-639-181	Sequence 181, App
139	7	1.5	566	9	US-10-173-706-56	Sequence 56, Appl	212	6	1.3	55	10	US-09-864-761-38777	Sequence 38777, A
140	7	1.5	566	9	US-10-175-738-56	Sequence 56, Appl	213	6	1.3	57	10	US-09-864-761-47578	Sequence 47578, A
141	7	1.5	566	9	US-10-175-752-56	Sequence 56, Appl	214	6	1.3	58	10	US-09-864-761-44799	Sequence 44799, A
142	7	1.5	566	9	US-10-176-482-56	Sequence 56, Appl	215	6	1.3	58	10	US-09-864-761-40506	Sequence 40506, A
143	7	1.5	566	9	US-10-176-757-56	Sequence 56, Appl	216	6	1.3	61	9	US-09-867-550-692	Sequence 692, App
144	7	1.5	566	9	US-10-176-913-56	Sequence 56, Appl	217	6	1.3	63	10	US-09-796-692-1165	Sequence 1165, Ap
145	7	1.5	566	9	US-10-180-552-56	Sequence 56, Appl	218	6	1.3	65	9	US-09-864-761-46591	Sequence 46591, A
146	7	1.5	566	9	US-10-180-557-56	Sequence 56, Appl	219	6	1.3	65	10	US-09-764-878-119	Sequence 119, App
147	7	1.5	566	10	US-09-989-722-41	Sequence 41, Appl	220	6	1.3	71	10	US-09-864-761-45224	Sequence 42524, A
148	7	1.5	566	10	US-09-989-722-41	Sequence 41, Appl	221	6	1.3	73	9	US-10-063-547-30	Sequence 30, Appl
149	7	1.5	566	10	US-09-989-727-41	Sequence 41, Appl	222	6	1.3	73	9	US-10-063-616-30	Sequence 30, Appl
150	7	1.5	566	10	US-09-989-727-41	Sequence 41, Appl	223	6	1.3	73	9	US-10-063-502-30	Sequence 30, Appl
151	7	1.5	566	10	US-09-989-731-41	Sequence 41, Appl	224	6	1.3	73	10	US-09-864-761-35230	Sequence 35230, A
152	7	1.5	566	10	US-09-989-733-41	Sequence 41, Appl	225	6	1.3	73	12	US-10-006-867-30	Sequence 30, Appl
153	7	1.5	566	10	US-09-991-073-41	Sequence 41, Appl	226	6	1.3	76	9	US-10-101-487-36	Sequence 36, Appl
154	7	1.5	566	10	US-09-990-442-41	Sequence 41, Appl	227	6	1.3	76	10	US-09-864-761-45186	Sequence 45186, A
155	7	1.5	566	10	US-09-991-163-41	Sequence 41, Appl	228	6	1.3	78	10	US-09-864-761-37973	Sequence 37973, A
156	7	1.5	566	10	US-09-993-604-41	Sequence 41, Appl	229	6	1.3	79	9	US-09-981-876-250	Sequence 250, App
157	7	1.5	566	10	US-09-990-456-41	Sequence 41, Appl	230	6	1.3	79	10	US-09-864-761-48376	Sequence 48376, A
158	7	1.5	566	12	US-10-052-586-56	Sequence 56, Appl	231	6	1.3	80	10	US-09-903-456-9	Sequence 9, Appl1
159	7	1.5	663	9	US-10-108-605-245	Sequence 245, App	232	6	1.3	83	10	US-09-864-761-48130	Sequence 48130, A
160	7	1.5	689	10	US-09-071-838-2	Sequence 2, Appl1	233	6	1.3	89	10	US-09-739-907-154	Sequence 154, App
161	7	1.5	714	10	US-09-978-243-3	Sequence 3, Appl1	234	6	1.3	89	10	US-09-864-761-42220	Sequence 42220, A
162	7	1.5	739	9	US-09-738-626-6773	Sequence 6773, Ap	235	6	1.3	90	10	US-09-739-907-94	Sequence 94, Appl
163	7	1.5	748	10	US-09-815-242-10278	Sequence 10278, A	236	6	1.3	91	10	US-09-764-877-1841	Sequence 1841, Ap
164	7	1.5	748	10	US-09-815-242-13880	Sequence 13880, A	237	6	1.3	91	12	US-10-001-879-160	Sequence 160, Appl
165	7	1.5	748	10	US-09-815-242-13880	Sequence 13880, A	238	6	1.3	102	10	US-09-745-763-6	Sequence 6, Appl1

239	6	1.3	105	10	US-09-925-301-1516	Sequence 1516, Ap	312	207	9	US-09-902-941-1913	Sequence 1913, Ap
240	6	1.3	108	9	US-09-950-933A-61	Sequence 61, Appl	313	207	9	US-09-738-626-4470	Sequence 5470, Ap
241	6	1.3	109	9	US-09-796-692-2460	Sequence 2460, Ap	314	207	9	US-09-849-626-1667	Sequence 1667, Ap
242	6	1.3	111	10	US-09-764-877-1774	Sequence 1774, Ap	315	207	9	US-09-849-626-1913	Sequence 1913, Ap
243	6	1.3	112	10	US-09-925-502-581	Sequence 581, App	316	208	10	US-09-214-881A-11	Sequence 11, Appl
244	6	1.3	114	9	US-09-738-626-3516	Sequence 3516, Ap	317	209	10	US-10-001-887-86	Sequence 86, Appl
245	6	1.3	114	9	US-09-925-300-1175	Sequence 1175, Ap	318	212	9	US-10-146-474-4	Sequence 4, Appl
246	6	1.3	115	10	US-10-079-623-351	Sequence 351, App	319	213	10	US-09-828-644-62	Sequence 62, Appl
247	6	1.3	115	10	US-09-903-456-23	Sequence 23, Appl	320	213	10	US-09-811-284-239	Sequence 239, App
248	6	1.3	116	10	US-09-764-869-689	Sequence 689, App	321	214	10	US-09-214-881A-3	Sequence 1, Appl
249	6	1.3	117	10	US-09-864-761-45592	Sequence 45592, A	322	214	10	US-09-214-881A-3	Sequence 3, Appl
250	6	1.3	122	10	US-09-729-674-6	Sequence 6, Appl	323	214	10	US-09-214-881A-4	Sequence 4, Appl
251	6	1.3	122	10	US-09-893-737-286	Sequence 286, App	324	214	10	US-09-214-881A-5	Sequence 5, Appl
252	6	1.3	123	9	US-09-726-258-17	Sequence 17, Appl	325	219	10	US-09-925-299-797	Sequence 797, App
253	6	1.3	124	9	US-10-051-767-8	Sequence 8, Appl	326	219	12	US-10-016-447-13	Sequence 13, Appl
254	6	1.3	124	9	US-10-031-767-14	Sequence 14, Appl	327	223	12	US-10-005-168-2	Sequence 2, Appl
255	6	1.3	126	9	US-09-738-626-4171	Sequence 4171, Ap	328	231	9	US-09-448-378-2	Sequence 2, Appl
256	6	1.3	127	9	US-09-864-761-38973	Sequence 38973, A	329	231	10	US-09-983-806-2	Sequence 2, Appl
257	6	1.3	128	10	US-09-925-300-1757	Sequence 1757, Ap	330	231	10	US-09-864-761-37211	Sequence 37211, A
258	6	1.3	130	10	US-09-741-669-294	Sequence 294, App	331	231	10	US-09-983-806-2	Sequence 2, Appl
259	6	1.3	130	10	US-09-912-020-318	Sequence 318, App	332	234	9	US-10-001-857-176	Sequence 176, App
260	6	1.3	130	10	US-09-864-761-45365	Sequence 45365, A	333	234	9	US-10-076-885-22	Sequence 22, Appl
261	6	1.3	130	10	US-09-815-242-10345	Sequence 10345, A	334	237	9	US-09-726-258-5	Sequence 25, Appl
262	6	1.3	130	9	US-09-815-242-14104	Sequence 14104, A	335	238	9	US-09-862-802-8	Sequence 8, Appl
263	6	1.3	140	9	US-09-879-461-4	Sequence 4, Appl	336	239	10	US-09-910-071-15	Sequence 15, Appl
264	6	1.3	140	10	US-09-764-864-858	Sequence 858, App	337	240	9	US-09-801-368-300	Sequence 300, App
265	6	1.3	146	10	US-09-940-497-3	Sequence 3, Appl	338	240	9	US-09-738-626-4338	Sequence 4338, App
266	6	1.3	146	10	US-09-969-834-3	Sequence 3, Appl	339	240	9	US-09-987-755-7	Sequence 7, Appl
267	6	1.3	153	10	US-09-764-887-206	Sequence 206, Appl	340	240	10	US-09-938-885A-3	Sequence 3, Appl
268	6	1.3	154	10	US-09-939-825-24	Sequence 24, Appl	341	240	10	US-09-801-368-342	Sequence 342, App
269	6	1.3	157	10	US-09-999-699-2	Sequence 2, Appl	342	246	10	US-09-925-302-511	Sequence 511, App
270	6	1.3	157	10	US-09-999-699-2	Sequence 6, Appl	343	247	9	US-09-908-193-42	Sequence 42, Appl
271	6	1.3	159	10	US-09-814-122-54	Sequence 54, Appl	344	248	9	US-09-908-193-41	Sequence 41, Appl
272	6	1.3	160	10	US-09-731-872-294	Sequence 294, App	345	249	9	US-09-908-193-46	Sequence 46, Appl
273	6	1.3	161	10	US-09-925-297-493	Sequence 493, App	346	249	9	US-10-213-700-4	Sequence 4, Appl
274	6	1.3	161	10	US-09-747-155-243	Sequence 243, App	347	251	9	US-10-213-700-3	Sequence 3, Appl
275	6	1.3	164	10	US-09-925-300-1010	Sequence 1010, Ap	348	253	10	US-09-305-856B-16	Sequence 16, Appl
276	6	1.3	169	10	US-09-764-846-207	Sequence 207, App	349	254	10	US-09-925-301-13408	Sequence 13408, Ap
277	6	1.3	170	10	US-09-976-165-22	Sequence 22, Appl	350	255	10	US-09-815-242-11399	Sequence 1399, A
278	6	1.3	174	10	US-09-903-456-40	Sequence 40, Appl	351	255	10	US-09-815-242-13599	Sequence 13599, A
279	6	1.3	175	9	US-10-050-786-11	Sequence 11, Appl	352	255	10	US-09-881-752A-106	Sequence 106, App
280	6	1.3	180	10	US-09-780-717-11	Sequence 11, Appl	353	256	10	US-09-965-529-25	Sequence 25, Appl
281	6	1.3	181	10	US-09-864-761-33314	Sequence 33314, A	354	258	9	US-09-815-242-10664	Sequence 10664, A
282	6	1.3	181	10	US-09-864-761-41658	Sequence 41658, A	355	260	9	US-09-738-626-65921	Sequence 5921, Ap
283	6	1.3	182	10	US-09-811-284-182	Sequence 182, App	356	261	10	US-09-925-301-901	Sequence 901, App
284	6	1.3	183	10	US-09-925-301-1230	Sequence 240, App	357	261	12	US-10-003-759-2	Sequence 175, App
285	6	1.3	183	9	US-10-028-072-240	Sequence 240, App	358	265	9	US-09-981-876-175	Sequence 9, Appl
286	6	1.3	189	9	US-10-121-049-240	Sequence 240, App	359	266	9	US-10-131-965-9	Sequence 9, Appl
287	6	1.3	189	9	US-10-123-904-240	Sequence 240, App	360	266	9	US-10-131-965-15	Sequence 15, Appl
288	6	1.3	189	9	US-10-140-470-240	Sequence 240, App	361	267	10	US-09-822-885-18	Sequence 8, Appl
289	6	1.3	189	10	US-09-864-761-34813	Sequence 34813, A	362	267	10	US-09-425-921-13	Sequence 13, Appl
290	6	1.3	189	10	US-09-867-550-1504	Sequence 1504, Ap	363	267	12	US-10-023-992-7	Sequence 7, Appl
291	6	1.3	189	10	US-09-975-999-2	Sequence 2, Appl	364	268	9	US-10-081-347-3	Sequence 33, Appl
292	6	1.3	189	10	US-09-975-374A-2	Sequence 2, Appl	365	268	10	US-09-284-663A-10	Sequence 10, Appl
293	6	1.3	191	10	US-09-764-870-362	Sequence 362, App	366	268	10	US-09-750-963-11	Sequence 11, Appl
294	6	1.3	191	10	US-09-925-300-1565	Sequence 1565, App	367	268	10	US-09-902-773A-7	Sequence 7, Appl
295	6	1.3	198	10	US-09-927-738-7	Sequence 7, Appl	368	268	10	US-09-251-263-12	Sequence 12, Appl
296	6	1.3	199	9	US-10-001-054-38	Sequence 38, Appl	369	268	10	US-09-796-338A-19	Sequence 19, Appl
297	6	1.3	199	9	US-10-028-072-32	Sequence 32, Appl	370	269	9	US-09-738-626-4086	Sequence 4086, Ap
298	6	1.3	199	9	US-10-121-049-92	Sequence 92, Appl	371	270	10	US-09-350-874-59	Sequence 59, Appl
299	6	1.3	199	9	US-10-123-904-92	Sequence 92, Appl	372	271	10	US-09-795-693-36	Sequence 36, Appl
300	6	1.3	199	9	US-10-140-470-92	Sequence 92, Appl	373	272	9	US-09-738-626-4596	Sequence 4596, Ap
301	6	1.3	200	8	US-08-424-550B-395	Sequence 395, App	374	272	9	US-09-738-626-6621	Sequence 6621, Ap
302	6	1.3	200	9	US-08-736-457-324	Sequence 324, App	375	273	9	US-10-218-381-1	Sequence 1, Appl
303	6	1.3	200	9	US-09-736-457-789	Sequence 789, App	376	277	10	US-09-755-456-7	Sequence 7, Appl
304	6	1.3	200	9	US-09-902-941-324	Sequence 324, App	377	280	8	US-08-910-386A-14	Sequence 14, Appl
305	6	1.3	200	9	US-09-902-941-789	Sequence 789, App	378	280	9	US-09-712-363-283	Sequence 283, App
306	6	1.3	200	9	US-09-849-626-324	Sequence 324, App	379	280	10	US-09-983-306-1	Sequence 1, Appl
307	6	1.3	201	10	US-09-849-626-789	Sequence 789, App	380	280	10	US-09-903-456-28	Sequence 28, Appl
308	6	1.3	201	10	US-09-214-881A-10	Sequence 10, Appl	381	282	9	US-09-738-626-511	Sequence 511, Ap
309	6	1.3	206	10	US-09-214-881A-9	Sequence 9, Appl	382	286	10	US-09-903-456-59	Sequence 59, Appl
310	6	1.3	207	9	US-09-736-457-1667	Sequence 1667, Ap	383	289	10	US-09-903-456-21	Sequence 21, Appl
311	6	1.3	207	9	US-09-902-941-1667	Sequence 1667, Ap	384	289	10	US-09-903-456-34	Sequence 34, Appl



385	6	1.3	291	10	US-09-903-456-36	Sequence 36, Appl	458	6	1.3	330	10	US-09-886-055-211	Sequence 211, App
386	6	1.3	293	10	US-09-903-456-19	Sequence 19, Appl	459	6	1.3	332	9	US-09-510-332-101	Sequence 101, App
387	6	1.3	294	10	US-09-927-738-9	Sequence 9, Appl1	460	6	1.3	333	9	US-09-976-059-2	Sequence 2, Appl1
388	6	1.3	295	9	US-09-978-295A-415	Sequence 415, App	461	6	1.3	333	9	US-09-738-626-5173	Sequence 5173, Ap
389	6	1.3	295	9	US-09-981-876-179	Sequence 179, App	462	6	1.3	333	9	US-10-108-105-6	Sequence 6, Appl1
390	6	1.3	295	9	US-09-978-697-415	Sequence 415, App	463	6	1.3	333	10	US-09-771-209-71	Sequence 71, Appl
391	6	1.3	295	9	US-09-999-852A-415	Sequence 415, App	464	6	1.3	333	10	US-09-825-882-2	Sequence 2, Appl1
392	6	1.3	295	9	US-09-978-159-415	Sequence 415, App	465	6	1.3	334	10	US-09-925-299-932	Sequence 1103, Ap
393	6	1.3	295	9	US-09-981-876-237	Sequence 237, App	466	6	1.3	335	9	US-09-738-626-6886	Sequence 6886, Ap
396	6	1.3	301	10	US-09-903-456-33	Sequence 33, Appl	468	6	1.3	336	9	US-09-976-059-6	Sequence 6, Appl1
397	6	1.3	302	10	US-09-815-242-11153	Sequence 11153, A	469	6	1.3	341	9	US-10-017-736-14	Sequence 14, Appl1
398	6	1.3	303	10	US-09-815-242-10382	Sequence 10382, A	470	6	1.3	344	10	US-09-740-288A-8	Sequence 8, Appl1
399	6	1.3	306	9	US-09-738-626-6782	Sequence 6782, Ap	471	6	1.3	346	10	US-09-815-242-11567	Sequence 11567, A
400	6	1.3	310	9	US-09-510-332-164	Sequence 164, App	472	6	1.3	347	10	US-09-903-456-14	Sequence 14, Appl
401	6	1.3	310	10	US-09-305-856B-14	Sequence 14, Appl	473	6	1.3	348	10	US-09-939-521-11	Sequence 11, Appl
402	6	1.3	311	10	US-09-898-416-9	Sequence 9, Appl1	474	6	1.3	352	9	US-10-017-736-13	Sequence 13, Appl
403	6	1.3	313	9	US-09-992-598-97	Sequence 97, Appl	475	6	1.3	352	9	US-09-799-777-45	Sequence 45, Appl
404	6	1.3	313	9	US-09-981-876-233	Sequence 233, App	476	6	1.3	359	10	US-09-925-301-1068	Sequence 1068, Ap
405	6	1.3	313	9	US-09-983-293A-97	Sequence 97, Appl	477	6	1.3	360	10	US-09-925-301-1068	Sequence 20, Appl
406	6	1.3	313	9	US-09-989-735-97	Sequence 97, Appl	478	6	1.3	365	10	US-09-993-348-20	Sequence 20, Appl
407	6	1.3	313	9	US-09-990-444-97	Sequence 97, Appl	479	6	1.3	365	10	US-09-908-711-100	Sequence 100, App
408	6	1.3	313	9	US-09-989-730-97	Sequence 97, Appl	480	6	1.3	366	9	US-09-900-425A-4	Sequence 4, Appl1
409	6	1.3	313	9	US-09-990-436-97	Sequence 97, Appl	481	6	1.3	366	10	US-10-076-785-24	Sequence 24, Appl1
410	6	1.3	313	9	US-09-991-181-97	Sequence 97, Appl	482	6	1.3	366	10	US-09-815-242-12345	Sequence 12345, A
411	6	1.3	313	9	US-09-993-687-97	Sequence 97, Appl	483	6	1.3	366	10	US-09-815-242-12768	Sequence 12768, A
412	6	1.3	313	9	US-09-989-734-97	Sequence 97, Appl	484	6	1.3	380	8	US-09-888-615-71	Sequence 71, Appl
413	6	1.3	313	9	US-10-028-072-398	Sequence 398, App	485	6	1.3	380	9	US-10-017-736-12	Sequence 12, Appl
414	6	1.3	313	9	US-09-997-653-97	Sequence 97, Appl	486	6	1.3	381	9	US-09-738-626-3549	Sequence 3549, Ap
415	6	1.3	313	9	US-09-993-667-97	Sequence 97, Appl	487	6	1.3	386	10	US-09-839-497A-5	Sequence 5, Appl1
416	6	1.3	313	9	US-10-121-049-398	Sequence 398, App	488	6	1.3	391	9	US-09-815-242-5379	Sequence 5379, Ap
417	6	1.3	313	9	US-10-123-904-398	Sequence 398, App	489	6	1.3	393	9	US-10-017-736-11	Sequence 11, Appl
418	6	1.3	313	9	US-09-989-722-97	Sequence 97, Appl	490	6	1.3	393	10	US-09-815-242-11285	Sequence 11285, A
419	6	1.3	313	10	US-09-989-723-97	Sequence 97, Appl	491	6	1.3	397	10	US-09-815-242-11925	Sequence 11925, A
420	6	1.3	313	10	US-09-989-723-97	Sequence 97, Appl	492	6	1.3	397	10	US-09-815-242-12002	Sequence 12002, A
421	6	1.3	313	10	US-09-989-729-97	Sequence 97, Appl	493	6	1.3	398	9	US-09-741-669-430	Sequence 430, App
422	6	1.3	313	10	US-09-989-727-97	Sequence 97, Appl	494	6	1.3	409	10	US-10-017-736-2	Sequence 2, Appl1
423	6	1.3	313	10	US-09-989-731-97	Sequence 97, Appl	495	6	1.3	409	10	US-09-778-927A-80	Sequence 80, Appl
424	6	1.3	313	10	US-09-989-732-97	Sequence 97, Appl	496	6	1.3	411	10	US-09-749-728B-25	Sequence 25, Appl
425	6	1.3	313	10	US-09-991-073-97	Sequence 97, Appl	497	6	1.3	414	10	US-09-925-331-16	Sequence 16, Appl
426	6	1.3	313	10	US-09-990-442-97	Sequence 97, Appl	498	6	1.3	414	10	US-09-927-738-8	Sequence 8, Appl1
427	6	1.3	313	10	US-09-991-163-97	Sequence 97, Appl	499	6	1.3	416	10	US-09-741-669-384	Sequence 384, App
428	6	1.3	313	10	US-09-993-604-97	Sequence 97, Appl	500	6	1.3	418	9	US-09-992-331-17	Sequence 17, Appl
429	6	1.3	313	10	US-09-990-456-97	Sequence 97, Appl	501	6	1.3	422	10	US-09-767-041-53	Sequence 53, Appl
430	6	1.3	314	10	US-09-989-721-97	Sequence 97, Appl	502	6	1.3	425	10	US-09-809-545A-53	Sequence 53, Appl
431	6	1.3	314	9	US-09-510-332-33	Sequence 33, Appl	503	6	1.3	426	9	US-09-815-242-10467	Sequence 10467, A
432	6	1.3	314	9	US-09-886-055-371	Sequence 371, App	504	6	1.3	428	9	US-09-738-626-6036	Sequence 6036, Ap
433	6	1.3	315	9	US-10-104-966-5	Sequence 5, Appl1	505	6	1.3	428	9	US-10-066-500-17	Sequence 17, Appl
434	6	1.3	315	10	US-09-815-242-11775	Sequence 11775, A	506	6	1.3	428	9	US-10-028-072-292	Sequence 292, App
435	6	1.3	315	10	US-09-898-416-8	Sequence 8, Appl1	507	6	1.3	428	9	US-09-992-331-15	Sequence 15, Appl
436	6	1.3	315	10	US-09-983-306-3	Sequence 3, Appl1	508	6	1.3	428	9	US-09-992-331-16	Sequence 16, Appl
437	6	1.3	315	10	US-09-983-306-6	Sequence 6, Appl1	509	6	1.3	428	9	US-10-121-049-292	Sequence 292, App
438	6	1.3	315	10	US-09-929-955-5	Sequence 5, Appl1	510	6	1.3	428	9	US-10-123-904-292	Sequence 292, App
439	6	1.3	316	10	US-09-886-055-197	Sequence 197, App	511	6	1.3	428	9	US-10-140-470-292	Sequence 292, App
440	6	1.3	317	10	US-09-305-856B-12	Sequence 12, Appl	512	6	1.3	429	10	US-09-347-331-10	Sequence 10, Appl
441	6	1.3	317	10	US-09-903-456-13	Sequence 13, Appl	513	6	1.3	431	10	US-09-815-242-113506	Sequence 13506, A
442	6	1.3	318	9	US-09-903-170C-3	Sequence 3, Appl1	514	6	1.3	434	10	US-09-764-86A-820	Sequence 820, App
443	6	1.3	318	9	US-09-957-156-2	Sequence 2, Appl1	515	6	1.3	434	10	US-09-778-927A-79	Sequence 79, Appl
444	6	1.3	318	10	US-09-739-907-76	Sequence 76, Appl	516	6	1.3	436	9	US-09-738-626-5887	Sequence 5887, Ap
445	6	1.3	318	10	US-09-903-180B-3	Sequence 3, Appl1	517	6	1.3	440	10	US-09-815-242-11399	Sequence 11399, A
446	6	1.3	318	10	US-09-903-187A-3	Sequence 3, Appl1	518	6	1.3	441	9	US-09-738-626-5934	Sequence 5934, Ap
447	6	1.3	318	10	US-09-903-171A-3	Sequence 3, Appl1	519	6	1.3	441	10	US-09-950-071-2	Sequence 2, Appl1
448	6	1.3	318	10	US-09-903-188A-3	Sequence 3, Appl1	520	6	1.3	442	9	US-10-063-547-76	Sequence 76, Appl
449	6	1.3	318	10	US-09-903-323A-3	Sequence 3, Appl1	521	6	1.3	442	9	US-10-063-616-76	Sequence 76, Appl
450	6	1.3	318	10	US-09-903-325A-3	Sequence 3, Appl1	522	6	1.3	442	9	US-10-063-502-76	Sequence 76, Appl
451	6	1.3	318	10	US-09-903-456-25	Sequence 25, Appl	523	6	1.3	442	12	US-10-006-867-76	Sequence 867, Ap
452	6	1.3	319	9	US-09-966-459A-11	Sequence 11, Appl	524	6	1.3	450	10	US-09-912-020-398	Sequence 398, App
453	6	1.3	321	10	US-09-815-242-4990	Sequence 4990, Ap	525	6	1.3	451	9	US-09-910-186A-12	Sequence 12, Appl
454	6	1.3	321	10	US-09-815-242-10635	Sequence 10635, A	526	6	1.3	451	10	US-09-764-86A-1124	Sequence 1124, Ap
455	6	1.3	323	9	US-09-791-932-119	Sequence 119, App	527	6	1.3	456	10	US-09-795-699-2	Sequence 2, Appl1
456	6	1.3	324	9	US-09-738-626-6547	Sequence 6547, Ap	528	6	1.3	457	10	US-09-052-753-1	Sequence 1, Appl1
457	6	1.3	329	10	US-09-925-301-904	Sequence 904, App	529	6	1.3	457	10	US-09-052-753-5	Sequence 5, Appl1
							530	6	1.3	463	9	US-10-155-613-1	Sequence 1, Appl1



531	6	1.3	466	10	US-09-778-927A-78	Sequence 78, Appl	604	6	1.3	604	9	US-09-738-626-3886	Sequence 3886, Ap
532	6	1.3	466	10	US-09-778-927A-81	Sequence 81, Appl	605	6	1.3	604	10	US-09-920-581-3	Sequence 3, Appl
533	6	1.3	468	10	US-09-860-232A-8	Sequence 8, Appl	606	6	1.3	608	9	US-09-738-626-5305	Sequence 5305, Ap
534	6	1.3	471	9	US-09-900-425A-5	Sequence 5, Appl	607	6	1.3	608	10	US-09-950-294-4	Sequence 4, Appl
535	6	1.3	473	10	US-09-826-752-14	Sequence 14, Appl	608	6	1.3	614	9	US-09-331-631A-21	Sequence 21, Appl
536	6	1.3	476	10	US-09-925-300-1582	Sequence 1582, Ap	609	6	1.3	619	10	US-09-815-242-11085	Sequence 11085, A
537	6	1.3	493	10	US-09-981-947A-10	Sequence 10, Appl	610	6	1.3	621	10	US-09-799-777-77	Sequence 77, Appl
538	6	1.3	493	10	US-09-815-242-5498	Sequence 5498, Ap	611	6	1.3	632	10	US-09-929-955-39	Sequence 29, Appl
539	6	1.3	494	10	US-09-815-242-12410	Sequence 12410, A	612	6	1.3	633	12	US-10-001-882-5	Sequence 5, Appl
540	6	1.3	500	10	US-09-323-998D-59	Sequence 59, Appl	613	6	1.3	650	9	US-09-738-626-4040	Sequence 4040, Ap
541	6	1.3	501	10	US-09-934-868-56	Sequence 56, Appl	614	6	1.3	656	9	US-09-991-496-6	Sequence 6, Appl
542	6	1.3	504	9	US-09-903-376-2	Sequence 2, Appl	615	6	1.3	661	10	US-09-874-923-6	Sequence 10, Appl
543	6	1.3	505	9	US-09-742-580-8	Sequence 8, Appl	616	6	1.3	666	10	US-10-114-893-10	Sequence 10, Appl
544	6	1.3	505	9	US-09-742-581-8	Sequence 8, Appl	617	6	1.3	662	9	US-10-047-542-80	Sequence 80, Appl
545	6	1.3	505	10	US-09-742-582-8	Sequence 8, Appl	618	6	1.3	680	10	US-09-996-184-10	Sequence 10, Appl
546	6	1.3	506	10	US-09-801-368-90	Sequence 90, Appl	619	6	1.3	686	9	US-09-881-239-3	Sequence 3, Appl
547	6	1.3	508	10	US-09-016-159-5	Sequence 5, Appl	620	6	1.3	686	10	US-09-929-955-17	Sequence 17, Appl
548	6	1.3	511	9	US-09-738-626-4770	Sequence 4770, Ap	621	6	1.3	686	10	US-09-929-955-31	Sequence 31, Appl
549	6	1.3	520	9	US-09-738-626-4493	Sequence 4493, Ap	622	6	1.3	686	10	US-09-929-955-32	Sequence 32, Appl
550	6	1.3	521	10	US-09-795-483-2	Sequence 2, Appl	623	6	1.3	686	10	US-09-929-955-43	Sequence 43, Appl
551	6	1.3	527	9	US-09-738-626-4803	Sequence 4803, Ap	624	6	1.3	686	10	US-09-929-955-44	Sequence 44, Appl
552	6	1.3	529	9	US-09-258-031B-16	Sequence 16, Appl	625	6	1.3	686	10	US-09-929-955-45	Sequence 45, Appl
553	6	1.3	529	9	US-09-258-031B-20	Sequence 20, Appl	626	6	1.3	686	10	US-09-929-955-46	Sequence 46, Appl
554	6	1.3	530	10	US-09-835-082-2	Sequence 2, Appl	627	6	1.3	686	10	US-09-929-955-47	Sequence 47, Appl
555	6	1.3	530	10	US-09-835-082-4	Sequence 4, Appl	628	6	1.3	686	10	US-09-929-955-48	Sequence 48, Appl
556	6	1.3	533	10	US-09-431-226-5	Sequence 5, Appl	629	6	1.3	686	10	US-09-929-955-49	Sequence 49, Appl
557	6	1.3	539	9	US-10-028-072-140	Sequence 140, App	630	6	1.3	686	10	US-09-881-654-2	Sequence 2, Appl
558	6	1.3	539	9	US-10-121-049-140	Sequence 140, App	631	6	1.3	697	10	US-09-881-752A-76	Sequence 76, Appl
559	6	1.3	539	9	US-10-123-904-140	Sequence 140, App	632	6	1.3	702	10	US-09-815-242-10117	Sequence 10117, A
560	6	1.3	539	9	US-10-140-470-140	Sequence 140, App	633	6	1.3	702	10	US-09-815-242-14005	Sequence 14005, A
561	6	1.3	542	9	US-10-028-072-188	Sequence 188, App	634	6	1.3	705	9	US-09-854-133-186	Sequence 186, App
562	6	1.3	542	9	US-10-174-5590-398	Sequence 398, App	635	6	1.3	705	10	US-09-925-302-531	Sequence 531, App
563	6	1.3	542	9	US-10-175-758-398	Sequence 398, App	636	6	1.3	705	10	US-09-738-973-186	Sequence 186, App
564	6	1.3	542	9	US-10-175-757-398	Sequence 398, App	637	6	1.3	709	9	US-10-118-328-4	Sequence 4, Appl
565	6	1.3	542	9	US-10-121-049-188	Sequence 188, App	638	6	1.3	717	10	US-09-817-913-9	Sequence 9, Appl
566	6	1.3	542	9	US-10-123-904-188	Sequence 188, App	639	6	1.3	717	10	US-09-817-958-9	Sequence 9, Appl
567	6	1.3	542	9	US-10-140-470-188	Sequence 140, App	640	6	1.3	724	9	US-09-908-193-25	Sequence 25, Appl
568	6	1.3	542	9	US-10-173-706-398	Sequence 398, App	641	6	1.3	728	9	US-09-881-228-1	Sequence 1, Appl
569	6	1.3	542	9	US-10-175-738-398	Sequence 398, App	642	6	1.3	729	9	US-10-108-603-251	Sequence 251, App
570	6	1.3	542	9	US-10-175-752-398	Sequence 398, App	643	6	1.3	731	9	US-09-361-630-3	Sequence 3, Appl
571	6	1.3	542	9	US-10-176-482-398	Sequence 398, App	644	6	1.3	731	9	US-10-086-464-17	Sequence 17, Appl
572	6	1.3	542	9	US-10-176-757-398	Sequence 398, App	645	6	1.3	734	9	US-09-738-626-4227	Sequence 4227, Ap
573	6	1.3	542	9	US-10-176-913-398	Sequence 398, App	646	6	1.3	746	9	US-09-981-353-30	Sequence 30, Appl
574	6	1.3	542	9	US-10-180-552-398	Sequence 398, App	647	6	1.3	746	9	US-09-982-107-4	Sequence 4, Appl
575	6	1.3	542	9	US-10-180-557-398	Sequence 398, App	648	6	1.3	747	9	US-09-974-143A-2	Sequence 2, Appl
576	6	1.3	542	12	US-10-052-586-398	Sequence 398, App	649	6	1.3	747	10	US-09-431-226-7	Sequence 7, Appl
577	6	1.3	544	10	US-09-967-624-6	Sequence 6, Appl	650	6	1.3	749	10	US-09-828-313-37	Sequence 27, Appl
578	6	1.3	546	10	US-09-746-359A-37	Sequence 37, Appl	651	6	1.3	757	10	US-09-928-175-34	Sequence 24, Appl
579	6	1.3	548	10	US-09-869-877-4	Sequence 4, Appl	652	6	1.3	763	9	US-10-081-408-2	Sequence 2, Appl
580	6	1.3	548	10	US-09-732-350-4	Sequence 4, Appl	653	6	1.3	764	9	US-09-981-353-59	Sequence 59, Appl
581	6	1.3	553	10	US-09-746-359A-11	Sequence 11, Appl	654	6	1.3	764	9	US-09-989-915-134	Sequence 124, Appl
582	6	1.3	553	10	US-09-949-192-7	Sequence 7, Appl	655	6	1.3	764	10	US-09-818-247-1	Sequence 1, Appl
583	6	1.3	560	9	US-09-912-672A-5	Sequence 5, Appl	656	6	1.3	769	9	US-09-815-242-10575	Sequence 10575, A
584	6	1.3	561	10	US-09-764-864-1539	Sequence 1539, Ap	657	6	1.3	772	9	US-09-738-626-6661	Sequence 6661, Ap
585	6	1.3	566	10	US-09-925-299-813	Sequence 813, App	658	6	1.3	772	10	US-09-935-799A-2	Sequence 2, Appl
586	6	1.3	571	10	US-09-815-242-10619	Sequence 10619, A	659	6	1.3	772	10	US-09-935-799A-5	Sequence 5, Appl
587	6	1.3	571	10	US-09-764-864-1028	Sequence 1028, Ap	660	6	1.3	775	9	US-09-992-598-326	Sequence 326, App
588	6	1.3	574	9	US-09-912-672A-2	Sequence 2, Appl	661	6	1.3	775	9	US-09-989-293A-326	Sequence 326, App
589	6	1.3	574	9	US-10-063-547-164	Sequence 164, App	662	6	1.3	775	9	US-09-989-735-336	Sequence 326, App
590	6	1.3	574	9	US-10-063-616-164	Sequence 164, App	663	6	1.3	775	9	US-09-990-444-336	Sequence 326, App
591	6	1.3	574	9	US-10-063-502-164	Sequence 164, App	664	6	1.3	775	9	US-09-989-730-336	Sequence 326, App
592	6	1.3	574	10	US-09-728-911-25	Sequence 25, Appl	665	6	1.3	775	9	US-09-990-438-326	Sequence 326, App
593	6	1.3	574	10	US-09-748-657-2	Sequence 2, Appl	666	6	1.3	775	9	US-09-991-181-336	Sequence 326, App
594	6	1.3	574	10	US-09-870-574-4	Sequence 4, Appl	667	6	1.3	775	9	US-09-993-687-336	Sequence 326, App
595	6	1.3	574	12	US-10-006-867-164	Sequence 164, App	668	6	1.3	775	9	US-09-993-734-336	Sequence 326, App
596	6	1.3	578	10	US-09-052-753-7	Sequence 7, Appl	669	6	1.3	775	9	US-09-997-653-336	Sequence 326, App
597	6	1.3	578	10	US-09-815-242-12501	Sequence 12501, A	670	6	1.3	775	9	US-09-993-667-336	Sequence 326, App
598	6	1.3	584	9	US-10-050-786-7	Sequence 7, Appl	671	6	1.3	775	10	US-09-989-723-326	Sequence 326, App
599	6	1.3	590	9	US-09-331-631A-8	Sequence 8, Appl	672	6	1.3	775	10	US-09-988-723-326	Sequence 326, App
600	6	1.3	599	10	US-09-952-013A-4	Sequence 4, Appl	673	6	1.3	775	10	US-09-989-279-326	Sequence 326, App
601	6	1.3	600	10	US-09-801-368-158	Sequence 158, App	674	6	1.3	775	10	US-09-989-727-326	Sequence 326, App
602	6	1.3	602	9	US-10-047-542-51	Sequence 51, Appl	675	6	1.3	775	10	US-09-989-731-326	Sequence 326, App
603	6	1.3	604	9	US-10-045-170A-1	Sequence 1, Appl	676	6	1.3	775	10	US-09-989-732-326	Sequence 326, App

677	6	1.3	775	10	US-09-991-073-326	Sequence 326, App	750	6	1.3	1332	9	US-10-041-856-5	Sequence 5, Appli
678	6	1.3	775	10	US-09-990-442-326	Sequence 326, App	751	6	1.3	1332	10	US-09-982-091A-4	Sequence 4, Appli
679	6	1.3	775	10	US-09-991-163-326	Sequence 326, App	752	6	1.3	1420	10	US-09-801-368-356	Sequence 336, App
680	6	1.3	775	10	US-09-993-604-326	Sequence 326, App	753	6	1.3	1427	9	US-09-991-496-97	Sequence 97, Appli
681	6	1.3	775	10	US-09-990-456-326	Sequence 326, App	754	6	1.3	1427	10	US-09-874-923-97	Sequence 97, Appli
682	6	1.3	775	10	US-09-989-721-326	Sequence 326, App	755	6	1.3	1464	9	US-09-842-777-10	Sequence 10, Appli
683	6	1.3	781	10	US-09-766-387-2	Sequence 2, Appli	756	6	1.3	1504	9	US-09-932-145-7	Sequence 7, Appli
684	6	1.3	781	10	US-09-766-387-4	Sequence 2, Appli	757	6	1.3	1507	9	US-10-017-216-6	Sequence 6, Appli
685	6	1.3	782	10	US-09-815-242-4893	Sequence 4893, Ap	758	6	1.3	1602	10	US-09-778-927A-59	Sequence 59, Appli
686	6	1.3	790	10	US-09-966-147-9	Sequence 9, Appli	759	6	1.3	1641	9	US-10-017-216-5	Sequence 5, Appli
687	6	1.3	794	10	US-09-789-404-2	Sequence 2, Appli	760	6	1.3	1641	9	US-09-991-466-96	Sequence 96, Appli
688	6	1.3	796	10	US-09-738-626-3660	Sequence 3660, Ap	761	6	1.3	1641	10	US-09-874-923-96	Sequence 96, Appli
689	6	1.3	811	10	US-09-815-242-11875	Sequence 11875, A	762	6	1.3	1643	10	US-09-738-626-6894	Sequence 6894, Ap
690	6	1.3	816	9	US-09-796-720B-2	Sequence 2, Appli	763	6	1.3	1736	10	US-09-919-497-98	Sequence 98, Appli
691	6	1.3	819	9	US-09-796-720B-4	Sequence 4, Appli	764	6	1.3	1809	10	US-09-919-497-98	Sequence 98, Appli
692	6	1.3	824	9	US-09-932-145-3	Sequence 3, Appli	765	6	1.3	1924	9	US-09-822-635-2	Sequence 2, Appli
693	6	1.3	829	9	US-09-881-239-5	Sequence 5, Appli	766	6	1.3	1958	12	US-10-028-946-4	Sequence 4, Appli
694	6	1.3	833	9	US-09-470-276-54	Sequence 54, Appli	767	6	1.3	2012	9	US-09-808-602-68	Sequence 68, Appli
695	6	1.3	847	9	US-09-870-759-52	Sequence 52, Appli	768	6	1.3	2053	9	US-10-017-216-2	Sequence 2, Appli
696	6	1.3	849	10	US-09-752-639-152	Sequence 152, App	769	6	1.3	2054	12	US-10-028-946-2	Sequence 2, Appli
697	6	1.3	849	10	US-09-984-198-152	Sequence 152, App	770	6	1.3	2055	9	US-10-017-216-2	Sequence 4, Appli
698	6	1.3	866	10	US-09-841-132-189	Sequence 189, App	771	6	1.3	2201	9	US-10-085-476-2	Sequence 2, Appli
699	6	1.3	875	12	US-10-001-852-2	Sequence 2, Appli	772	6	1.3	2291	9	US-10-000-789-2	Sequence 2, Appli
700	6	1.3	880	10	US-09-841-132-175	Sequence 175, App	773	6	1.3	2894	10	US-09-941-611-23	Sequence 23, Appli
701	6	1.3	888	9	US-10-036-041-35	Sequence 35, Appli	774	6	1.3	3011	9	US-09-747-419-20	Sequence 20, Appli
702	6	1.3	888	9	US-10-028-072-544	Sequence 544, App	775	6	1.3	3011	9	US-10-104-966-1	Sequence 1, Appli
703	6	1.3	888	9	US-10-035-855-35	Sequence 35, Appli	776	6	1.3	3011	9	US-09-891-894-3	Sequence 3, Appli
704	6	1.3	888	9	US-10-121-049-544	Sequence 544, App	777	6	1.3	3011	10	US-09-742-659-4	Sequence 4, Appli
705	6	1.3	888	9	US-10-123-904-544	Sequence 544, App	778	6	1.3	3011	10	US-09-916-359-2	Sequence 2, Appli
706	6	1.3	888	9	US-10-140-470-544	Sequence 544, App	779	6	1.3	3011	10	US-09-238-076-20	Sequence 20, Appli
707	6	1.3	888	12	US-10-036-342-35	Sequence 35, Appli	780	6	1.3	3011	10	US-09-952-572-9	Sequence 9, Appli
708	6	1.3	890	9	US-10-060-425-8	Sequence 8, Appli	781	6	1.3	3011	10	US-09-929-955-1	Sequence 1, Appli
709	6	1.3	890	9	US-10-060-425-10	Sequence 10, Appli	782	6	1.3	3012	10	US-09-238-076-2	Sequence 2, Appli
710	6	1.3	890	12	US-10-001-852-6	Sequence 6, Appli	783	6	1.3	3472	9	US-10-027-806-4	Sequence 4, Appli
711	6	1.3	911	10	US-10-066-320-2	Sequence 2, Appli	784	6	1.3	3472	9	US-10-034-653-4	Sequence 4, Appli
712	6	1.3	913	10	US-09-223-490-4	Sequence 4, Appli	785	6	1.3	4342	10	US-09-815-242-5107	Sequence 5107, Ap
713	6	1.3	918	12	US-10-001-852-7	Sequence 7, Appli	786	6	1.3	5701	10	US-09-864-761-37319	Sequence 37319, A
714	6	1.3	927	10	US-09-804-474A-4	Sequence 4, Appli	787	6	1.3	26926	6	US-09-759-508B-2	Sequence 28, Appli
715	6	1.3	950	10	US-09-321-987B-4	Sequence 4, Appli	788	5	1.1	6	1	US-08-809-423A-28	Sequence 12, Appli
716	6	1.3	961	10	US-09-220-091-3	Sequence 3, Appli	789	5	1.1	6	10	US-09-821-380-12	Sequence 12, Appli
717	6	1.3	962	9	US-09-738-626-3800	Sequence 3800, Ap	790	5	1.1	7	10	US-09-821-380-9	Sequence 9, Appli
718	6	1.3	967	12	US-10-105-929-2	Sequence 2, Appli	791	5	1.1	7	10	US-09-821-380-11	Sequence 11, Appli
719	6	1.3	968	9	US-10-163-316-7	Sequence 7, Appli	792	5	1.1	8	1	US-08-841-656A-7	Sequence 7, Appli
720	6	1.3	981	10	US-09-991-258-13	Sequence 13, Appli	793	5	1.1	8	10	US-09-821-380-10	Sequence 10, Appli
721	6	1.3	1002	10	US-09-988-117-3	Sequence 3, Appli	794	5	1.1	8	12	US-10-007-761-13	Sequence 13, Appli
722	6	1.3	1002	10	US-09-812-471-3	Sequence 3, Appli	795	5	1.1	9	8	US-08-871-076-4	Sequence 4, Appli
723	6	1.3	1002	10	US-09-812-633-3	Sequence 3, Appli	796	5	1.1	9	8	US-08-854-825-53	Sequence 53, Appli
724	6	1.3	1005	9	US-09-988-117-1	Sequence 1, Appli	797	5	1.1	9	9	US-09-277-074-14	Sequence 14, Appli
725	6	1.3	1005	10	US-09-812-471-1	Sequence 1, Appli	798	5	1.1	9	10	US-09-872-832-12	Sequence 12, Appli
726	6	1.3	1005	10	US-08-834-705-18	Sequence 18, Appli	799	5	1.1	9	10	US-09-791-378-541	Sequence 541, App
727	6	1.3	1019	1	US-10-027-591-2	Sequence 2, Appli	800	5	1.1	10	10	US-09-767-460-47	Sequence 47, App
728	6	1.3	1044	9	US-09-866-582-37	Sequence 37, Appli	801	5	1.1	10	10	US-09-821-380-3	Sequence 3, Appli
729	6	1.3	1047	10	US-09-147-052-4	Sequence 4, Appli	802	5	1.1	10	10	US-09-821-380-4	Sequence 4, Appli
730	6	1.3	1086	10	US-09-866-510-14	Sequence 14, Appli	803	5	1.1	10	10	US-09-791-378-635	Sequence 635, App
731	6	1.3	1090	9	US-09-932-145-4	Sequence 4, Appli	804	5	1.1	11	1	US-08-809-423A-5	Sequence 5, Appli
732	6	1.3	1099	10	US-09-881-654-4	Sequence 4, Appli	805	5	1.1	11	10	US-09-873-676-23	Sequence 23, Appli
733	6	1.3	1106	9	US-09-955-363-2	Sequence 2, Appli	806	5	1.1	12	10	US-09-734-520-75	Sequence 75, Appli
734	6	1.3	1106	10	US-09-866-510-16	Sequence 16, Appli	807	5	1.1	12	10	US-09-734-520-76	Sequence 76, Appli
735	6	1.3	1106	10	US-09-866-510-18	Sequence 18, Appli	808	5	1.1	12	12	US-10-012-034A-75	Sequence 75, Appli
736	6	1.3	1106	10	US-09-866-510-20	Sequence 20, Appli	809	5	1.1	12	12	US-10-012-034A-76	Sequence 76, Appli
737	6	1.3	1106	10	US-09-866-510-22	Sequence 22, Appli	810	5	1.1	13	9	US-09-851-138-156	Sequence 156, App
738	6	1.3	1106	10	US-09-866-510-22	Sequence 22, Appli	811	5	1.1	13	10	US-09-919-124-30	Sequence 30, Appli
739	6	1.3	1114	9	US-10-001-873-50	Sequence 50, Appli	812	5	1.1	13	10	US-09-821-380-2	Sequence 2, Appli
740	6	1.3	1163	10	US-09-893-348-18	Sequence 18, Appli	813	5	1.1	13	10	US-09-981-194-23	Sequence 23, Appli
741	6	1.3	1213	10	US-09-969-515-8	Sequence 8, Appli	814	5	1.1	13	10	US-09-943-692-29	Sequence 29, Appli
742	6	1.3	1224	10	US-09-969-515-2	Sequence 2, Appli	815	5	1.1	15	9	US-10-002-344A-208	Sequence 208, App
743	6	1.3	1229	10	US-09-815-242-5818	Sequence 5818, Ap	816	5	1.1	15	9	US-09-774-639-124	Sequence 124, App
744	6	1.3	1229	10	US-09-815-242-12946	Sequence 12946, A	817	5	1.1	16	9	US-09-759-130B-194	Sequence 194, App
745	6	1.3	1246	10	US-09-741-669-349	Sequence 349, App	818	5	1.1	16	10	US-09-818-954A-28	Sequence 28, Appli
746	6	1.3	1251	10	US-09-778-927A-58	Sequence 58, Appli	819	5	1.1	17	9	US-09-826-290-253	Sequence 253, App
747	6	1.3	1286	9	US-10-017-216-7	Sequence 7, Appli	820	5	1.1	17	9	US-09-826-290-255	Sequence 255, App
748	6	1.3	1332	9	US-10-041-856-3	Sequence 3, Appli	821	5	1.1	17	9	US-09-826-290-258	Sequence 258, App
749	6	1.3	1332	9	US-10-041-856-4	Sequence 4, Appli	822	5	1.1	17	9	US-09-826-290-448	Sequence 448, App

823	5	1.1	17	9	US-09-916-494A-26	Sequence 26, App1	896	5	1.1	26	9	US-09-349-755-7	Sequence 7, App1
824	5	1.1	17	9	US-09-759-130B-174	Sequence 174, App	897	5	1.1	26	9	US-09-349-755-14	Sequence 14, App1
825	5	1.1	17	9	US-09-759-130B-354	Sequence 354, App	898	5	1.1	26	9	US-09-166-334-7	Sequence 7, App1
826	5	1.1	17	10	US-09-864-761-46718	Sequence 46718, A	899	5	1.1	26	9	US-09-166-334-14	Sequence 14, App1
827	5	1.1	17	10	US-09-821-380-1	Sequence 1, App1	900	5	1.1	26	10	US-09-030-619-123	Sequence 123, App
828	5	1.1	17	10	US-09-798-116-12	Sequence 12, App1	901	5	1.1	26	10	US-09-864-761-35286	Sequence 35286, A
829	5	1.1	17	10	US-09-798-116-23	Sequence 23, App1	902	5	1.1	26	10	US-09-350-206-7	Sequence 7, App1
830	5	1.1	17	10	US-09-791-378-4	Sequence 4, App1	903	5	1.1	26	10	US-09-350-206-14	Sequence 14, App1
831	5	1.1	17	10	US-09-791-378-79	Sequence 79, App1	904	5	1.1	26	10	US-09-952-782-4	Sequence 4, App1
832	5	1.1	17	10	US-09-791-378-111	Sequence 111, App	905	5	1.1	26	10	US-09-952-782-23	Sequence 23, App1
833	5	1.1	17	10	US-09-791-378-114	Sequence 114, App	906	5	1.1	27	9	US-10-117-057-8	Sequence 8, App1
834	5	1.1	17	10	US-09-791-378-159	Sequence 159, App	907	5	1.1	27	9	US-09-966-459A-17	Sequence 17, App1
835	5	1.1	17	10	US-09-791-378-162	Sequence 162, App	908	5	1.1	27	9	US-09-983-802-921	Sequence 373, App
836	5	1.1	17	10	US-09-791-378-166	Sequence 166, App	909	5	1.1	27	10	US-09-864-761-35226	Sequence 35226, A
837	5	1.1	17	10	US-09-791-378-229	Sequence 229, App	910	5	1.1	27	10	US-09-864-761-44867	Sequence 44867, A
838	5	1.1	17	10	US-09-791-378-237	Sequence 237, App	911	5	1.1	27	10	US-09-952-782-22	Sequence 22, App1
839	5	1.1	17	10	US-09-791-378-240	Sequence 240, App	912	5	1.1	28	9	US-10-012-896-921	Sequence 921, App
840	5	1.1	17	10	US-09-791-378-274	Sequence 274, App	913	5	1.1	28	9	US-09-895-793-921	Sequence 921, App
841	5	1.1	17	10	US-09-791-378-307	Sequence 307, App	914	5	1.1	28	9	US-09-895-814-921	Sequence 921, App
842	5	1.1	17	10	US-09-791-378-308	Sequence 308, App	915	5	1.1	28	10	US-09-759-143-921	Sequence 921, App
843	5	1.1	17	10	US-09-791-378-340	Sequence 340, App	916	5	1.1	28	10	US-09-864-761-33674	Sequence 33674, A
844	5	1.1	17	10	US-09-791-378-621	Sequence 621, App	917	5	1.1	28	10	US-09-864-761-33674	Sequence 33674, A
845	5	1.1	17	10	US-09-791-378-647	Sequence 647, App	918	5	1.1	28	10	US-09-864-761-37130	Sequence 37130, A
846	5	1.1	18	9	US-10-074-956-20	Sequence 20, App1	919	5	1.1	28	10	US-09-780-669-921	Sequence 921, App
847	5	1.1	18	9	US-09-162-095-4	Sequence 4, App1	920	5	1.1	28	10	US-09-822-827-921	Sequence 921, App
848	5	1.1	18	10	US-09-864-761-48194	Sequence 48194, A	921	5	1.1	28	10	US-09-822-827-921	Sequence 921, App
849	5	1.1	18	10	US-09-952-782-31	Sequence 31, App1	922	5	1.1	28	10	US-09-952-782-21	Sequence 21, App1
850	5	1.1	19	9	US-10-078-547-22	Sequence 22, App1	923	5	1.1	29	9	US-08-424-550B-486	Sequence 486, App
851	5	1.1	19	10	US-09-864-761-42761	Sequence 42761, A	924	5	1.1	29	9	US-09-996-140-230	Sequence 230, App
852	5	1.1	19	10	US-09-952-782-30	Sequence 30, App1	925	5	1.1	29	9	US-09-996-140-231	Sequence 231, App
853	5	1.1	19	10	US-09-734-520-15	Sequence 15, App1	926	5	1.1	29	9	US-09-996-140-232	Sequence 232, App
854	5	1.1	19	12	US-10-012-034A-15	Sequence 15, App1	927	5	1.1	29	9	US-10-223-047-16	Sequence 16, App1
855	5	1.1	20	9	US-10-007-270-29	Sequence 29, App1	928	5	1.1	29	10	US-09-864-761-44607	Sequence 44607, A
856	5	1.1	20	9	US-09-915-676-9	Sequence 9, App1	929	5	1.1	29	10	US-09-864-761-47524	Sequence 47524, A
857	5	1.1	20	10	US-09-812-528-18	Sequence 18, App1	930	5	1.1	30	9	US-09-917-340-30	Sequence 30, App1
858	5	1.1	20	10	US-09-466-320-6	Sequence 6, App1	931	5	1.1	30	10	US-09-852-797-79	Sequence 79, App1
859	5	1.1	20	10	US-09-952-782-29	Sequence 29, App1	932	5	1.1	30	10	US-09-864-761-34474	Sequence 34474, A
860	5	1.1	21	8	US-08-871-076-6	Sequence 6, App1	933	5	1.1	30	10	US-09-864-761-37450	Sequence 37450, A
861	5	1.1	21	9	US-09-909-460-15	Sequence 15, App1	934	5	1.1	30	10	US-09-864-761-40773	Sequence 40773, A
862	5	1.1	21	10	US-09-952-782-28	Sequence 28, App1	935	5	1.1	30	10	US-09-137-531-12	Sequence 12, App1
863	5	1.1	22	10	US-09-833-740-9	Sequence 9, App1	936	5	1.1	30	10	US-09-853-161-79	Sequence 19, App1
864	5	1.1	22	10	US-09-864-761-38066	Sequence 38066, A	937	5	1.1	31	9	US-09-852-659A-79	Sequence 79, App1
865	5	1.1	22	10	US-09-864-761-38136	Sequence 38136, A	938	5	1.1	31	9	US-10-042-431-20	Sequence 20, App1
866	5	1.1	22	10	US-09-875-494-19	Sequence 19, App1	939	5	1.1	31	10	US-09-759-130B-390	Sequence 390, App1
867	5	1.1	22	10	US-09-952-782-27	Sequence 27, App1	940	5	1.1	31	10	US-09-986-676A-112	Sequence 112, App1
868	5	1.1	23	9	US-09-924-400-133	Sequence 133, App1	941	5	1.1	32	10	US-09-764-877-1239	Sequence 1239, App
869	5	1.1	23	9	US-09-969-920-171	Sequence 171, App	942	5	1.1	32	10	US-09-864-761-34495	Sequence 34495, A
870	5	1.1	23	10	US-09-746-919-41	Sequence 41, App1	943	5	1.1	32	10	US-09-864-761-41454	Sequence 41454, A
871	5	1.1	23	10	US-09-746-919-42	Sequence 42, App1	944	5	1.1	32	10	US-09-864-761-43444	Sequence 43444, A
872	5	1.1	23	10	US-09-832-336-6	Sequence 6, App1	945	5	1.1	32	10	US-09-864-761-45034	Sequence 45034, A
873	5	1.1	23	10	US-09-810-936-133	Sequence 133, App	946	5	1.1	32	10	US-09-864-761-45811	Sequence 45811, A
874	5	1.1	23	10	US-09-952-782-26	Sequence 26, App1	947	5	1.1	32	10	US-09-864-761-46737	Sequence 46737, A
875	5	1.1	23	10	US-09-429-755-133	Sequence 133, App	948	5	1.1	32	10	US-09-864-761-47019	Sequence 47019, A
876	5	1.1	24	9	US-10-002-344A-218	Sequence 218, App	949	5	1.1	33	9	US-09-737-379-7	Sequence 7, App1
877	5	1.1	24	9	US-09-759-130B-159	Sequence 159, App	950	5	1.1	33	10	US-09-983-802-236	Sequence 236, App
878	5	1.1	24	10	US-09-739-907-142	Sequence 142, App	951	5	1.1	33	10	US-09-864-761-44510	Sequence 44510, A
879	5	1.1	24	10	US-09-864-761-38460	Sequence 38460, A	952	5	1.1	33	10	US-09-864-761-44885	Sequence 44885, A
880	5	1.1	24	10	US-09-864-761-39114	Sequence 39114, A	953	5	1.1	33	10	US-09-764-869-1122	Sequence 1122, App
881	5	1.1	24	10	US-09-864-761-42315	Sequence 42315, A	954	5	1.1	34	9	US-09-759-130B-74	Sequence 74, App1
882	5	1.1	24	10	US-09-952-782-25	Sequence 25, App1	955	5	1.1	34	10	US-09-864-761-38085	Sequence 38085, A
883	5	1.1	25	8	US-08-424-550B-180	Sequence 180, App	956	5	1.1	34	10	US-09-864-761-41080	Sequence 41080, A
884	5	1.1	25	9	US-09-731-449-10	Sequence 10, App1	957	5	1.1	34	10	US-09-864-761-44095	Sequence 44095, A
885	5	1.1	25	9	US-09-989-920-214	Sequence 214, App	958	5	1.1	34	10	US-09-864-761-47878	Sequence 47878, A
886	5	1.1	25	9	US-09-759-130B-77	Sequence 77, App1	959	5	1.1	34	10	US-09-938-719-13	Sequence 13, App1
887	5	1.1	25	10	US-09-729-472-3	Sequence 3, App1	960	5	1.1	34	10	US-09-938-703-13	Sequence 13, App1
888	5	1.1	25	10	US-09-821-984-31	Sequence 31, App1	961	5	1.1	35	9	US-09-774-639-223	Sequence 223, App
889	5	1.1	25	10	US-09-738-970-3	Sequence 3, App1	962	5	1.1	35	10	US-09-864-761-35975	Sequence 35975, A
890	5	1.1	25	10	US-09-864-761-36489	Sequence 36489, A	963	5	1.1	35	10	US-09-864-761-40336	Sequence 40336, A
891	5	1.1	25	10	US-09-864-761-40555	Sequence 40555, A	964	5	1.1	35	10	US-09-864-761-42339	Sequence 42339, A
892	5	1.1	25	10	US-09-770-693-17	Sequence 17, App1	965	5	1.1	35	10	US-09-864-761-46401	Sequence 46401, A
893	5	1.1	25	10	US-09-729-499-3	Sequence 3, App1	966	5	1.1	35	10	US-09-864-761-47358	Sequence 47358, A
894	5	1.1	25	10	US-09-952-782-22	Sequence 22, App1	967	5	1.1	36	9	US-09-852-797-82	Sequence 82, App1
895	5	1.1	25	10	US-09-226-248B-22	Sequence 22, App1	968	5	1.1	36	9	US-09-983-802-182	Sequence 182, App

```
969 5 1.1 36 10 US-09-864-761-37154 Sequence 37154, A
970 -5 1.1 36 10 US-09-864-761-37406 Sequence 37406, A
971 5 1.1 36 10 US-09-864-761-40594 Sequence 40594, A
972 -5 1.1 36 10 US-09-864-761-40624 Sequence 40624, A
973 5 1.1 36 10 US-09-864-761-41813 Sequence 41813, A
974 5 1.1 36 10 US-09-864-761-43021 Sequence 43021, A
975 5 1.1 36 10 US-09-853-161-82 Sequence 82, App1
976 5 1.1 36 10 US-09-852-659A-82 Sequence 336, App
977 5 1.1 37 10 US-09-764-860-336 Sequence 179, App
978 5 1.1 37 10 US-08-809-423A-14 Sequence 14, App
979 5 1.1 37 10 US-10-042-141-179 Sequence 179, App
980 5 1.1 37 9 US-10-042-141-184 Sequence 184, App
981 5 1.1 37 10 US-09-726-643-179 Sequence 179, App
982 5 1.1 37 10 US-09-726-643-184 Sequence 184, App
983 5 1.1 37 10 US-09-864-761-34464 Sequence 34464, A
984 5 1.1 37 10 US-09-864-761-34924 Sequence 34924, A
985 5 1.1 37 10 US-09-864-761-37682 Sequence 37682, A
986 5 1.1 37 10 US-09-864-761-47013 Sequence 47013, A
987 5 1.1 37 10 US-09-764-869-930 Sequence 930, App1
988 5 1.1 37 10 US-09-772-719-6 Sequence 6, App1
989 5 1.1 38 9 US-09-836-392-30 Sequence 30, App1
990 5 1.1 38 9 US-10-042-141-167 Sequence 167, App
991 5 1.1 38 10 US-09-726-643-167 Sequence 167, App
992 5 1.1 38 10 US-09-864-761-39173 Sequence 39173, A
993 5 1.1 38 10 US-09-864-761-48776 Sequence 48776, A
994 5 1.1 39 9 US-09-774-639-274 Sequence 274, App
995 5 1.1 39 9 US-09-774-639-351 Sequence 351, App
996 5 1.1 39 10 US-09-864-761-33798 Sequence 33798, A
997 5 1.1 39 10 US-09-864-761-38733 Sequence 38733, A
998 5 1.1 39 10 US-09-864-761-38738 Sequence 38738, A
999 5 1.1 39 10 US-09-864-761-39865 Sequence 39865, A
1000 5 1.1 39 10 US-09-864-761-40246 Sequence 40246, A
```

## ALIGNMENTS

```
RESULT 1
US-09-754-949-4
; Sequence 4, Application US/09754949
; Patent No. US20020015939A1
GENERAL INFORMATION:
; APPLICANT: MCCARTHY, JUSTIN
; APPLICANT: CORDELL, BARBARA
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INHIBITORS OF
; FILE REFERENCE: SCIOS 012A
; CURRENT APPLICATION NUMBER: US/09/754,949
; CURRENT FILING DATE: 2001-01-04
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 467
; TYPE: PRN
; ORGANISM: Homo Sapien
US-09-754-949-4
```

Query Match 78.4%; Score 366; DB 10; Length 467;

Best Local Similarity 99.8%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;

```
Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MTELPAPLSTYFONAKMSDNHLSNTVRSQNDNREHNDNRSLGHEPLSNRPOGNSR 60
DB 1 MTELPAPLSTYFONAKMSDNHLSNTVRSQNDNREHNDNRSLGHEPLSNRPOGNSR 60
QY 61 QVVEODEEDELTLKYGAKHVIMLEVPVTLCAVNVVATIKSVFYTRKDGOLITPFT 120
DB 61 QVVEODEEDELTLKYGAKHVIMLEVPVTLCAVNVVATIKSVFYTRKDGOLITPFT 120
QY 121 DPEVVGORALHSLINAIMISVIVVMTILLVLYKRYCVIHAWLIISSLLFFPSFI 180
DB 121 DPEVVGORALHSLINAIMISVIVVMTILLVLYKRYCVIHAWLIISSLLFFPSFI 180
```

```
QY 181 YLGEVFTYVNAVADYITVALLINMLGVVGMISIMKGPLRLQOAYLIMISALMALVEIKY 240
DB 181 YLGEVFTYVNAVADYITVALLINMLGVVGMISIMKGPLRLQOAYLIMISALMALVEIKY 240
QY 241 LPEVTAMLLIIVISVYDLVAVLCPKGPLRMLVETAOERNETLFPALISSTWMLVMAE 300
DB 241 LPEVTAMLLIIVISVYDLVAVLCPKGPLRMLVETAOERNETLFPALISSTWMLVMAE 300
QY 301 GDPBAORRVKSNKYNAESTERESQDTVAENDDGFSEEMEAORDSHLGHPRSTPESRAA 360
DB 301 GDPBAORRVKSNKYNAESTERESQDTVAENDDGFSEEMEAORDSHLGHPRSTPESRAA 360
QY 361 VOELSSSILAGEDPEERGVKLGIDFIYSVLVGKASATASGDMNTTIACFAVAILGLCL 420
DB 361 VOELSSSILAGEDPEERGVKLGIDFIYSVLVGKASATASGDMNTTIACFAVAILGLCL 420
```

```
RESULT 2
US-09-878-454A-3
; Sequence 3, Application US/09878454A
; Patent No. US20020064828A1
GENERAL INFORMATION:
; APPLICANT: Monteliro, et al.
; TITLE OF INVENTION: Method of controlling the binding of Calmyrin to Presenilin
; FILE REFERENCE: 4115-161
; CURRENT APPLICATION NUMBER: US/09/878,454A
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/210,939
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 467
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-878-454A-3
```

Query Match 78.4%; Score 366; DB 10; Length 467;

Best Local Similarity 99.8%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;

```
Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MTELPAPLSTYFONAKMSDNHLSNTVRSQNDNREHNDNRSLGHEPLSNRPOGNSR 60
DB 1 MTELPAPLSTYFONAKMSDNHLSNTVRSQNDNREHNDNRSLGHEPLSNRPOGNSR 60
QY 61 QVVEODEEDELTLKYGAKHVIMLEVPVTLCAVNVVATIKSVFYTRKDGOLITPFT 120
DB 61 QVVEODEEDELTLKYGAKHVIMLEVPVTLCAVNVVATIKSVFYTRKDGOLITPFT 120
QY 121 DPEVVGORALHSLINAIMISVIVVMTILLVLYKRYCVIHAWLIISSLLFFPSFI 180
DB 121 DPEVVGORALHSLINAIMISVIVVMTILLVLYKRYCVIHAWLIISSLLFFPSFI 180
QY 181 YLGEVFTYVNAVADYITVALLINMLGVVGMISIMKGPLRLQOAYLIMISALMALVEIKY 240
DB 181 YLGEVFTYVNAVADYITVALLINMLGVVGMISIMKGPLRLQOAYLIMISALMALVEIKY 240
QY 241 LPEVTAMLLIIVISVYDLVAVLCPKGPLRMLVETAOERNETLFPALISSTWMLVMAE 300
DB 241 LPEVTAMLLIIVISVYDLVAVLCPKGPLRMLVETAOERNETLFPALISSTWMLVMAE 300
QY 301 GDPBAORRVKSNKYNAESTERESQDTVAENDDGFSEEMEAORDSHLGHPRSTPESRAA 360
DB 301 GDPBAORRVKSNKYNAESTERESQDTVAENDDGFSEEMEAORDSHLGHPRSTPESRAA 360
QY 361 VOELSSSILAGEDPEERGVKLGIDFIYSVLVGKASATASGDMNTTIACFAVAILGLCL 420
DB 361 VOELSSSILAGEDPEERGVKLGIDFIYSVLVGKASATASGDMNTTIACFAVAILGLCL 420
```

Oy 421 TLLLAIFKKALPALPISITFGVFAFDYLVOPFMOQLAHOFYI 467  
|  
Db 421 TLLLAIFKKALPALPISITFGVFAFDYLVOPFMOQLAHOFYI 467

## RESULT 3

US-09-895-035-14  
; Sequence 14, Application US/09895035  
; Patent No. US20020082211A1  
; GENERAL INFORMATION:

; APPLICANT: Patterson, Chandra  
; APPLICANT: Kaser, Lynn E.  
; APPLICANT: Kaser, Matthew R.  
; TITLE OF INVENTION: HUMAN PRESENTILIN VARIANT  
; FILE REFERENCE: PC-0047 CIP  
; CURRENT APPLICATION NUMBER: US/09/895,035  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: 09/116,640  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PERL Program  
; SEQ ID NO 14  
; LENGTH: 467  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incyte ID No. US20020082211A1 g1709856  
US-09-895-035-14

Query Match 78.4%; Score 366; DB 10; Length 467;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 MTELPAPLSYFQNAQMSQEDNHLSTVRSQNDNRERQEHNDRLSLGHPPLSGRPGNSR 60  
|  
Db 1 MTELPAPLSYFQNAQMSQEDNHLSTVRSQNDNRERQEHNDRLSLGHPPLSGRPGNSR 60  
Oy 61 OVEDEDEDEDELTKYGAHHVIMLFVPTLCMVVAVATIKSVFTRKDGOLITPTPE 120  
|  
Db 61 OVEDEDEDEDELTKYGAHHVIMLFVPTLCMVVAVATIKSVFTRKDGOLITPTPE 120  
Oy 121 DTEYGORALHSILNAAMISIVVMTLLVLYKRCYKVIHAWLIISLLFFFSFI 180  
|  
Db 121 DTEYGORALHSILNAAMISIVVMTLLVLYKRCYKVIHAWLIISLLFFFSFI 180  
Oy 181 YLGEVFKTYNNAVADYITVALLIMNLGVGMISIHMKGPLRLQOAYLIMISALMALVFIKY 240  
|  
Db 181 YLGEVFKTYNNAVADYITVALLIMNLGVGMISIHMKGPLRLQOAYLIMISALMALVFIKY 240  
Oy 241 LPEMTAMILLAVISYVDLVAVICPKGPLMLVETAOERNETLFPALITYSSTVWMLVNAE 300  
|  
Db 241 LPEMTAMILLAVISYVDLVAVICPKGPLMLVETAOERNETLFPALITYSSTVWMLVNAE 300  
Oy 301 GDPQAQRVSKSKYNAESTERESQDTVAENDDGFSEEMEAQRSHLGRHSTPESRAA 360  
|  
Db 301 GDPQAQRVSKSKYNAESTERESQDTVAENDDGFSEEMEAQRSHLGRHSTPESRAA 360  
Oy 361 VOELSSSILAGDPPERGVKGIGDFIFYSVLVGRASATASGDMWTTTACFVAAILIGLCL 420  
|  
Db 361 VOELSSSILAGDPPERGVKGIGDFIFYSVLVGRASATASGDMWTTTACFVAAILIGLCL 420  
Oy 421 TLLLAIFKKALPALPISITFGVFAFDYLVOPFMOQLAHOFYI 467  
|  
Db 421 TLLLAIFKKALPALPISITFGVFAFDYLVOPFMOQLAHOFYI 467

## RESULT 4

US-10-071-900-1  
; Sequence 1, Application US/10071900  
; Patent No. US20020127541A1  
; GENERAL INFORMATION:  
; APPLICANT: St. George-Hyslop, Peter H.

; APPLICANT: Rommens, Johanna  
; APPLICANT: Fraser, Paul E.  
; TITLE OF INVENTION: Alzheimer's Related Proteins and Methods  
; FILE REFERENCE: 1034/1F810-US1  
; CURRENT APPLICATION NUMBER: US/10/071,900  
; CURRENT FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: US/09/727,725  
; PRIOR FILING DATE: 1999-01-08  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 467  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-071-900-1

Query Match 78.4%; Score 366; DB 12; Length 467;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 MTELPAPLSYFQNAQMSQEDNHLSTVRSQNDNRERQEHNDRLSLGHPPLSGRPGNSR 60  
|  
Db 1 MTELPAPLSYFQNAQMSQEDNHLSTVRSQNDNRERQEHNDRLSLGHPPLSGRPGNSR 60  
Oy 61 OVEDEDEDEDELTKYGAHHVIMLFVPTLCMVVAVATIKSVFTRKDGOLITPTPE 120  
|  
Db 61 OVEDEDEDEDELTKYGAHHVIMLFVPTLCMVVAVATIKSVFTRKDGOLITPTPE 120  
Oy 121 DTEYGORALHSILNAAMISIVVMTLLVLYKRCYKVIHAWLIISLLFFFSFI 180  
|  
Db 121 DTEYGORALHSILNAAMISIVVMTLLVLYKRCYKVIHAWLIISLLFFFSFI 180  
Oy 181 YLGEVFKTYNNAVADYITVALLIMNLGVGMISIHMKGPLRLQOAYLIMISALMALVFIKY 240  
|  
Db 181 YLGEVFKTYNNAVADYITVALLIMNLGVGMISIHMKGPLRLQOAYLIMISALMALVFIKY 240  
Oy 241 LPEMTAMILLAVISYVDLVAVICPKGPLMLVETAOERNETLFPALITYSSTVWMLVNAE 300  
|  
Db 241 LPEMTAMILLAVISYVDLVAVICPKGPLMLVETAOERNETLFPALITYSSTVWMLVNAE 300  
Oy 301 GDPQAQRVSKSKYNAESTERESQDTVAENDDGFSEEMEAQRSHLGRHSTPESRAA 360  
|  
Db 301 GDPQAQRVSKSKYNAESTERESQDTVAENDDGFSEEMEAQRSHLGRHSTPESRAA 360  
Oy 361 VOELSSSILAGDPPERGVKGIGDFIFYSVLVGRASATASGDMWTTTACFVAAILIGLCL 420  
|  
Db 361 VOELSSSILAGDPPERGVKGIGDFIFYSVLVGRASATASGDMWTTTACFVAAILIGLCL 420  
Oy 421 TLLLAIFKKALPALPISITFGVFAFDYLVOPFMOQLAHOFYI 467  
|  
Db 421 TLLLAIFKKALPALPISITFGVFAFDYLVOPFMOQLAHOFYI 467

## RESULT 5

US-09-895-035-12  
; Sequence 12, Application US/09895035  
; Patent No. US20020082211A1  
; GENERAL INFORMATION:

; APPLICANT: Patterson, Chandra  
; APPLICANT: Kaser, Lynn E.  
; APPLICANT: Kaser, Matthew R.  
; TITLE OF INVENTION: HUMAN PRESENTILIN VARIANT  
; FILE REFERENCE: PC-0047 CIP  
; CURRENT APPLICATION NUMBER: US/09/895,035  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: 09/116,640  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PERL Program  
; SEQ ID NO 12  
; LENGTH: 463  
; TYPE: PRT

ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incycle ID No. US20020082211A1 g1244638  
US-09-895-035-12

Query Match 72.2% Score 337; DB 10; Length 463;  
Best Local Similarity 99.8% Pred. No. 2e-304;  
Matches 437; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

30 NDNREOEHNRRSLGHEPLISNGRPOGNSROYVODEDEDELTKGAKHVMLEFPV 89  
26 NDNREOEHNRRSLGHEPLISNGRPOGNSROYVODEDEDELTKGAKHVMLEFPV 85  
90 TLCAVVAVATIKSVSFYTRKDGOLYPTFETETVCGORALHSILNAIMISIVVMTL 149  
86 TLCAVVAVATIKSVSFYTRKDGOLYPTFETETVCGORALHSILNAIMISIVVMTL 145  
150 LVLLYKRCYVIAHMLIISLLIFFESFIYLGVEFKTVNAVDTITVALLINMGVVG 209  
146 LVLLYKRCYVIAHMLIISLLIFFESFIYLGVEFKTVNAVDTITVALLINMGVVG 205  
210 MISHMKEPLAQAYLIMISALMALVFIKYLEPWTAMLLAVISVYDLVAVLCPKGLR 269  
206 MISHMKEPLAQAYLIMISALMALVFIKYLEPWTAMLLAVISVYDLVAVLCPKGLR 265  
270 MLVETAOENETLFPALISSTWVWLVMAEGDEPAORRVKSKYNAESTERESODTVA 329  
266 MLVETAOENETLFPALISSTWVWLVMAEGDEPAORRVKSKYNAESTERESODTVA 325  
330 ENDDGGSEEAORDSLGHPRSTPESRAAVOELSSSILAGEDPERGVKLGJDFIY 389  
326 ENDDGGSEEAORDSLGHPRSTPESRAAVOELSSSILAGEDPERGVKLGJDFIY 385  
390 SVLVGKASATASGDMNTTACFVAITLIGLCLTLLLAIFKKALPALPISITFGVYFAT 449  
386 SVLVGKASATASGDMNTTACFVAITLIGLCLTLLLAIFKKALPALPISITFGVYFAT 445  
450 DYLVQPEMDQLAFHQFYI 467  
446 DYLVQPEMDQLAFHQFYI 463

RESULT 6  
US-09-785-474-2  
Sequence 2, Application US/09785474  
Patent No. US20010012626A1  
GENERAL INFORMATION:  
APPLICANT: TANZI, RUDOLPH  
WASCO, WILMA

TITLE OF INVENTION: Genetic Alterations Related To Familial  
Alzheimer's Disease  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVENUE, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/785,474  
FILING DATE: 20-Feb-2001  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/706,344  
FILING DATE: 30-AUG-1996  
APPLICATION NUMBER: 60/003,054

FILING DATE: 31-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: KIM, JUDITH U.  
REGISTRATION NUMBER: 40,679  
REFERENCE/DOCKET NUMBER: 0609.4180002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-785-474-2

Query Match 56.7% Score 265; DB 10; Length 467;  
Best Local Similarity 99.6% Pred. No. 1e-237;  
Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 MTELPAPLSYFONQMSDNHLSNTRVQNDNREOEHNDRSLGHEPLISNGRPOGNSR 60  
1 MTELPAPLSYFONQMSDNHLSNTRVQNDNREOEHNDRSLGHEPLISNGRPOGNSR 60  
61 QVDEDEDEDELTKYAKHVMLEFPVTLCAVVAVATIKSVSFYTRKDGOLYPTFTE 120  
61 QVDEDEDEDELTKYAKHVMLEFPVTLCAVVAVATIKSVSFYTRKDGOLYPTFTE 120  
121 DTEVCGORALHSILNAIMISIVVMTLLVLYKRCYVIAHMLIISLLIFFESFI 180  
121 DTEVCGORALHSILNAIMISIVVMTLLVLYKRCYVIAHMLIISLLIFFESFI 180  
181 YLGEVFKTVNAVDTITVALLINMGVVGMSIHKKGLRQOAYLIMISALMALVFIKY 240  
181 YLGEVFKTVNAVDTITVALLINMGVVGMSIHKKGLRQOAYLIMISALMALVFIKY 240  
241 LPEWTAMLLAVISVYDLVAVLCPKGLRMLVETAOENETLFPALISSTWVWLVMAE 300  
241 LPEWTAMLLAVISVYDLVAVLCPKGLRMLVETAOENETLFPALISSTWVWLVMAE 300  
301 GDPEAORRVKSKYNAESTERESODTVAENDDGGFSEEAORDSLGHPRSTPESRAA 360  
301 GDPEAORRVKSKYNAESTERESODTVAENDDGGFSEEAORDSLGHPRSTPESRAA 360  
361 VOELSSSILAGEDPERGVKLGJDFIYSLVGKASATASGDMNTTACFVAITLIGLCL 420  
361 VOELSSSILAGEDPERGVKLGJDFIYSLVGKASATASGDMNTTACFVAITLIGLCL 420  
421 TLLLAIFKKALPALPISITFGVYFATDYLVQPEMDQLAFHQFYI 467  
421 TLLLAIFKKALPALPISITFGVYFATDYLVQPEMDQLAFHQFYI 467

RESULT 7  
US-09-785-474-4  
Sequence 4, Application US/09785474  
Patent No. US20010012626A1  
GENERAL INFORMATION:  
APPLICANT: TANZI, RUDOLPH  
WASCO, WILMA

TITLE OF INVENTION: Genetic Alterations Related To Familial  
Alzheimer's Disease  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVENUE, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/785,474  
FILING DATE: 20-Feb-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/706,344  
FILING DATE: 30-AUG-1996  
APPLICATION NUMBER: 60/003,054  
FILING DATE: 31-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: KIM, JUDITH U.  
REGISTRATION NUMBER: 40,679  
REFERENCE/DOCKET NUMBER: 0609,4180002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-785-474-4

Query Match 43.7%; Score 204; DB 10; Length 467;  
Best Local Similarity 100.0%; Pred. No. 3.5e-181;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTELPAFLSYFQNAQSEDNHLSNTVRSQNDNRERQEHNDRLSLGHPEPLSNGRPOGNSR 60  
DB 1 MTELPAFLSYFQNAQSEDNHLSNTVRSQNDNRERQEHNDRLSLGHPEPLSNGRPOGNSR 60  
QY 61 QVDEDEDEDELTLLKYGAKHYIMLFVPTLCMVVAVATIKSVSYTRKDGOLITPPTFE 120  
DB 61 QVDEDEDEDELTLLKYGAKHYIMLFVPTLCMVVAVATIKSVSYTRKDGOLITPPTFE 120  
QY 121 DTEYVQGRALHSILNAAMISIVVMTILLVLYKRYCKVIHAWLIISLLFFFSFI 180  
DB 121 DTEYVQGRALHSILNAAMISIVVMTILLVLYKRYCKVIHAWLIISLLFFFSFI 180  
QY 181 YLGEVFKTYNVAVDYITVALLIWN 204  
DB 181 YLGEVFKTYNVAVDYITVALLIWN 204

RESULT 8  
US-09-785-474-28

; Sequence 28, Application US/09785474  
; Patent No. US20010012626A1  
; GENERAL INFORMATION:  
; APPLICANT: TANZI, RUDOLPH  
; MASCO, WILMA  
; TITLE OF INVENTION: Genetic Alterations Related To Familial  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVENUE, SUITE 600  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/785,474

FILING DATE: 20-Feb-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/706,344  
FILING DATE: 30-AUG-1996  
APPLICATION NUMBER: 60/003,054  
FILING DATE: 31-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: KIM, JUDITH U.  
REGISTRATION NUMBER: 40,679  
REFERENCE/DOCKET NUMBER: 0609,4180002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 28:  
US-09-785-474-28

Query Match 43.7%; Score 204; DB 10; Length 467;  
Best Local Similarity 100.0%; Pred. No. 3.5e-181;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTELPAFLSYFQNAQSEDNHLSNTVRSQNDNRERQEHNDRLSLGHPEPLSNGRPOGNSR 60  
DB 1 MTELPAFLSYFQNAQSEDNHLSNTVRSQNDNRERQEHNDRLSLGHPEPLSNGRPOGNSR 60  
QY 61 QVDEDEDEDELTLLKYGAKHYIMLFVPTLCMVVAVATIKSVSYTRKDGOLITPPTFE 120  
DB 61 QVDEDEDEDELTLLKYGAKHYIMLFVPTLCMVVAVATIKSVSYTRKDGOLITPPTFE 120  
QY 121 DTEYVQGRALHSILNAAMISIVVMTILLVLYKRYCKVIHAWLIISLLFFFSFI 180  
DB 121 DTEYVQGRALHSILNAAMISIVVMTILLVLYKRYCKVIHAWLIISLLFFFSFI 180  
QY 181 YLGEVFKTYNVAVDYITVALLIWN 204  
DB 181 YLGEVFKTYNVAVDYITVALLIWN 204

## RESULT 9

US-09-785-474-30  
; Sequence 30, Application US/09785474  
; Patent No. US20010012626A1  
; GENERAL INFORMATION:  
; APPLICANT: TANZI, RUDOLPH  
; MASCO, WILMA  
; TITLE OF INVENTION: Genetic Alterations Related To Familial  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVENUE, SUITE 600  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/785,474  
; FILING DATE: 20-Feb-2001  
; CLASSIFICATION: <unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/706,344  
; FILING DATE: 30-AUG-1996

APPLICATION NUMBER: 60/003,054  
FILING DATE: 31-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: KIM, JUDITH U.  
REGISTRATION NUMBER: 40,679  
REFERENCE/DOCKET NUMBER: 0609,4180002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 30:  
US-09-785-474-30

Query Match 43.7%; Score 204; DB 10; Length 467;  
Best Local Similarity 100.0%; Pred. No. 3.5e-181;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTELPAPLSTYFQNAQMSDNLSTNRQNDNRERQEHNDRLSLGHPPLSGRPGNSR 60  
|||||  
DB 1 MTELPAPLSTYFQNAQMSDNLSTNRQNDNRERQEHNDRLSLGHPPLSGRPGNSR 60  
QY 61 QVEODEEDELTLKYAKHVMLEFVPTLCMVVAVATIKSVSYTRKDGOLITPTE 120  
|||||  
DB 61 QVEODEEDELTLKYAKHVMLEFVPTLCMVVAVATIKSVSYTRKDGOLITPTE 120  
QY 121 DTEYVGRALSHILNAIMISIVYVMTLLVLYKRCYKVIHAWLIISLLLEFFSFI 180  
|||||  
DB 121 DTEYVGRALSHILNAIMISIVYVMTLLVLYKRCYKVIHAWLIISLLLEFFSFI 180  
QY 181 YLGEVFKTYNVAVDYITVALLIMN 204  
|||||  
DB 181 YLGEVFKTYNVAVDYITVALLIMN 204

## RESULT 10

US-09-785-474-32  
Sequence 32, Application US/09785474  
Patent No. US20010012626A1  
GENERAL INFORMATION:  
APPLICANT: TANZI, RUDOLPH  
MASCO, WILLA  
TITLE OF INVENTION: Genetic Alterations Related To Familial  
Alzheimer's Disease  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVENUE, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/785,474  
FILING DATE: 20-Feb-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/706,344  
FILING DATE: 30-AUG-1996  
APPLICATION NUMBER: 60/003,054  
FILING DATE: 31-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: KIM, JUDITH U.  
REGISTRATION NUMBER: 40,679

REFERENCE/DOCKET NUMBER: 0609,4180002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 32:  
US-09-785-474-32

Query Match 43.7%; Score 204; DB 10; Length 467;  
Best Local Similarity 100.0%; Pred. No. 3.5e-181;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTELPAPLSTYFQNAQMSDNLSTNRQNDNRERQEHNDRLSLGHPPLSGRPGNSR 60  
|||||  
DB 1 MTELPAPLSTYFQNAQMSDNLSTNRQNDNRERQEHNDRLSLGHPPLSGRPGNSR 60  
QY 61 QVEODEEDELTLKYAKHVMLEFVPTLCMVVAVATIKSVSYTRKDGOLITPTE 120  
|||||  
DB 61 QVEODEEDELTLKYAKHVMLEFVPTLCMVVAVATIKSVSYTRKDGOLITPTE 120  
QY 121 DTEYVGRALSHILNAIMISIVYVMTLLVLYKRCYKVIHAWLIISLLLEFFSFI 180  
|||||  
DB 121 DTEYVGRALSHILNAIMISIVYVMTLLVLYKRCYKVIHAWLIISLLLEFFSFI 180  
QY 181 YLGEVFKTYNVAVDYITVALLIMN 204  
|||||  
DB 181 YLGEVFKTYNVAVDYITVALLIMN 204

## RESULT 11

US-09-895-035-1  
Sequence 1, Application US/09895035  
Patent No. US2002008221A1  
GENERAL INFORMATION:  
APPLICANT: Patterson, Chandra  
Murry, Lynn E.  
APPLICANT: Kaser, Matthew R.  
TITLE OF INVENTION: HUMAN PRESENILIN VARIANT  
FILE REFERENCE: PC-0047 CIP  
CURRENT APPLICATION NUMBER: US/09/895,035  
PRIOR FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: 09/116,640  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PERL Program  
SEQ ID NO 1  
LENGTH: 180  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID NO. US2002008221A1 1353337CD1  
US-09-895-035-1

Query Match 28.3%; Score 132; DB 10; Length 180;  
Best Local Similarity 100.0%; Pred. No. 7.6e-115;  
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 NDNREORHNDRLSLGHPPLSGRPGNSRQVVEODEEDELTLKYAKHVMLEFVPV 89  
|||||  
DB 26 NDNREORHNDRLSLGHPPLSGRPGNSRQVVEODEEDELTLKYAKHVMLEFVPV 85  
QY 90 TLCMVVAVATIKSVSYTRKDGOLITPTEDETVGGRALSHILNAIMISIVYVMTL 149  
|||||  
DB 86 TLCMVVAVATIKSVSYTRKDGOLITPTEDETVGGRALSHILNAIMISIVYVMTL 145  
QY 150 LVVLYKRCYKVI 161  
|||||



Db 146 LVLYKRYCKV 157

RESULT 12

US-09-823-153-8

Sequence 8, Application US/09823153

Patent No. US20020025540A1

GENERAL INFORMATION:

APPLICANT: Bristol-Myers Squibb Company

APPLICANT: Roberts, Susan

APPLICANT: Pak, Roger

APPLICANT: Lewis, Martin

APPLICANT: Smith, David

APPLICANT: Hendrick, Joseph

APPLICANT: Vinitzky, Alexander

TITLE OF INVENTION: ISOLATION OF FUNCTIONALLY ACTIVE GAMMA-SECRETASE PROTEIN COMPLEX

TITLE OF INVENTION: METHODS FOR DETECTION OF ACTIVITY AND INHIBITORS THEREOF

FILE REFERENCE: D0004

CURRENT APPLICATION NUMBER: US/09/823,153

CURRENT FILING DATE: 2001-07-02

NUMBER OF SEQ ID NOS: 11

SOFTWARE: PatentIn version 3.0

SEQ ID NO 8

LENGTH: 354

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

NAME/KEY: misc\_feature

OTHER INFORMATION: Description of Artificial Sequence: PSI PEPTIDE

US-09-823-153-8

Query Match 27.4%; Score 128; DB 10; Length 354;

Best Local Similarity 100.0%; Pred. No. 7.3e-111;

Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 282 LFPALVYSTWVLYNNAEGPEAORVSKNSKYNAESTERESODTVVENDDGFSEEME 341

Db 227 LFPALVYSTWVLYNNAEGPEAORVSKNSKYNAESTERESODTVVENDDGFSEEME 286

QY 342 AORSHLGPRAHSTPESRAAVVELSSILAGDPPERGYKLGDFIFSVLVKASATAS 401

Db 287 AORSHLGPRAHSTPESRAAVVELSSILAGDPPERGYKLGDFIFSVLVKASATAS 346

QY 402 GDMNTTIA 409

Db 347 GDMNTTIA 354

RESULT 13

US-09-864-761-38056

Sequence 38056, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aeomica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263,6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

SEQ ID NO 38056

LENGTH: 65

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AF109907.1

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.7

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.8

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.7

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.2

OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.9

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2

OTHER INFORMATION: SWISSPROT HIT: P97887, EVALU8.00e-12

OTHER INFORMATION: EST\_HUMAN HIT: BE935454.1, EVALU8.1.00e-18

US-09-864-761-38056

Query Match 10.3%; Score 48; DB 10; Length 65;

Best Local Similarity 100.0%; Pred. No. 2e-37;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 IYPTFEDTETVYGORALHSILNAAIMISVIVMTLLVLVLYKRYCKV 161

Db 1 IYPTFEDTETVYGORALHSILNAAIMISVIVMTLLVLVLYKRYCKV 48

RESULT 14

US-09-925-299-1227

Sequence 1227, Application US/09925299

Patent No. US20020055627A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: PA102

CURRENT APPLICATION NUMBER: US/09/925,299

CURRENT FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: PCT/US00/05883

PRIOR FILING DATE: 2000-03-08

PRIOR APPLICATION NUMBER: 60/124,270

PRIOR FILING DATE: 1999-03-12

NUMBER OF SEQ ID NOS: 1556

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1227

```

: LENGTH: 101
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (32)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-1227

Query Match
Best Local Similarity 100.0%; Pred. No. 1,6e-33; Length 101;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 NETLEPALIYSTWVWLVNMAEGDEAQRVSKSKYNAESTER 322
|||||
Db 33 NETLEPALIYSTWVWLVNMAEGDEAQRVSKSKYNAESTER 76

RESULT 15
US-10-051-767-10
: Sequence 10, Application US/10051767
: Publication No. US20030022151A1
: GENERAL INFORMATION:
: APPLICANT: THINAKARAN, GOPAL
: TITLE OF INVENTION: FUNCTIONAL SCREENING
: FILE REFERENCE: ARCD:364US
: CURRENT APPLICATION NUMBER: US/10/051,767
: CURRENT FILING DATE: 2002-01-17
: PRIOR APPLICATION NUMBER: 60/262,353
: PRIOR FILING DATE: 2001-01-17
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 10
: LENGTH: 131
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Synthetic
: OTHER INFORMATION: Peptide
US-10-051-767-10

Query Match
Best Local Similarity 100.0%; Pred. No. 1,3e-18; Length 131;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 ISVYDLVAVLCPPKGPLRLMLVETAQERNE 280
|||||
Db 10 ISVYDLVAVLCPPKGPLRLMLVETAQERNE 37

RESULT 16
US-09-754-949-6
: Sequence 6, Application US/09754949
: Patent No. US20020015939A1
: GENERAL INFORMATION:
: APPLICANT: MCCARTHY, JUSTIN
: APPLICANT: CORDELL, BARBARA
: TITLE OF INVENTION: METHODS FOR IDENTIFYING INHIBITORS OF
: TITLE OF INVENTION: NEURONAL DEGENERATION
: FILE REFERENCE: SCIOS.012A
: CURRENT APPLICATION NUMBER: US/09/754,949
: CURRENT FILING DATE: 2001-01-04
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 6
: LENGTH: 448
: TYPE: PRT
: ORGANISM: Homo Sapien
US-09-754-949-6

Query Match
Best Local Similarity 100.0%; Pred. No. 4,2e-18; Length 448;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 253 ISVYDLVAVLCPPKGPLRLMLVETAQERNE 280
|||||
Db 259 ISVYDLVAVLCPPKGPLRLMLVETAQERNE 286

RESULT 17
US-09-878-454A-1
: Sequence 1, Application US/09878454A
: Patent No. US20020064828A1
: GENERAL INFORMATION:
: APPLICANT: Montelito, et al.
: TITLE OF INVENTION: Method of Controlling the Binding of Calmyrin to Presentin
: FILE REFERENCE: 4115-161
: CURRENT APPLICATION NUMBER: US/09/878,454A
: CURRENT FILING DATE: 2001-06-11
: PRIOR APPLICATION NUMBER: 60/210,939
: PRIOR FILING DATE: 2000-06-11
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 448
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-878-454A-1

Query Match
Best Local Similarity 100.0%; Pred. No. 4,2e-18; Length 448;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 ISVYDLVAVLCPPKGPLRLMLVETAQERNE 280
|||||
Db 259 ISVYDLVAVLCPPKGPLRLMLVETAQERNE 286

RESULT 18
US-10-071-900-2
: Sequence 2, Application US/10071900
: Patent No. US20020127541A1
: GENERAL INFORMATION:
: APPLICANT: St. George-Hyslop, Peter H.
: APPLICANT: Rommens, Johanna
: APPLICANT: Fraser, Paul E.
: TITLE OF INVENTION: Alzheimer's Related Proteins and Methods
: TITLE OF INVENTION: of Use
: FILE REFERENCE: 1034/1P810-US1
: CURRENT APPLICATION NUMBER: US/10/071,900
: CURRENT FILING DATE: 2002-02-08
: PRIOR APPLICATION NUMBER: US/09/227,725
: PRIOR FILING DATE: 1999-01-08
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 448
: TYPE: PRT
: ORGANISM: Homo Sapien
US-10-071-900-2

Query Match
Best Local Similarity 100.0%; Pred. No. 4,2e-18; Length 448;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 ISVYDLVAVLCPPKGPLRLMLVETAQERNE 280
|||||
Db 259 ISVYDLVAVLCPPKGPLRLMLVETAQERNE 286

RESULT 19
US-10-051-767-9
: Sequence 9, Application US/10051767
: Publication No. US20030022151A1
: GENERAL INFORMATION:
: APPLICANT: THINAKARAN, GOPAL
```

;; TITLE OF INVENTION: FUNCTIONAL SCREENING  
;; FILE REFERENCE: ARCD:364US  
;; CURRENT APPLICATION NUMBER: US/10/051,767  
;; CURRENT FILING DATE: 2002-01-17  
;; PRIOR APPLICATION NUMBER: 60/262,353  
;; PRIOR FILING DATE: 2001-01-17  
;; NUMBER OF SEQ ID NOS: 15  
;; SOFTWARE: Patentln Ver. 2.1  
;; SEQ ID NO 9  
;; LENGTH: 74  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
;; OTHER INFORMATION: Peptide  
US-10-051-767-9

Query Match 4.7%; Score 22; DB 9; Length 74;  
Best Local Similarity 100.0%; Pred. No. 2.8e-13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 375 EERGVKLGDFIFYSVLVGA 396  
Db 52 EERGVKLGDFIFYSVLVGA 73

RESULT 20  
US-10-051-767-15  
;; Sequence 15, Application US/10051767  
;; Publication No. US20030022151A1  
;; GENERAL INFORMATION:  
;; APPLICANT: THINAKARAN, GOPAL  
;; TITLE OF INVENTION: FUNCTIONAL SCREENING  
;; FILE REFERENCE: ARCD:364US  
;; CURRENT APPLICATION NUMBER: US/10/051,767  
;; CURRENT FILING DATE: 2002-01-17  
;; PRIOR APPLICATION NUMBER: 60/262,353  
;; PRIOR FILING DATE: 2001-01-17  
;; NUMBER OF SEQ ID NOS: 15  
;; SOFTWARE: Patentln Ver. 2.1  
;; SEQ ID NO 15  
;; LENGTH: 74  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
;; OTHER INFORMATION: Peptide  
US-10-051-767-15

Query Match 4.7%; Score 22; DB 9; Length 74;  
Best Local Similarity 100.0%; Pred. No. 2.8e-13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 375 EERGVKLGDFIFYSVLVGA 396  
Db 52 EERGVKLGDFIFYSVLVGA 73

RESULT 21  
US-10-051-767-11  
;; Sequence 11, Application US/10051767  
;; Publication No. US20030022151A1  
;; GENERAL INFORMATION:  
;; APPLICANT: THINAKARAN, GOPAL  
;; TITLE OF INVENTION: FUNCTIONAL SCREENING  
;; FILE REFERENCE: ARCD:364US  
;; CURRENT APPLICATION NUMBER: US/10/051,767  
;; CURRENT FILING DATE: 2002-01-17  
;; PRIOR APPLICATION NUMBER: 60/262,353  
;; PRIOR FILING DATE: 2001-01-17  
;; NUMBER OF SEQ ID NOS: 15  
;; SOFTWARE: Patentln Ver. 2.1  
;; SEQ ID NO 11

;; LENGTH: 129  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
;; OTHER INFORMATION: Peptide  
US-10-051-767-11

Query Match 4.7%; Score 22; DB 9; Length 129;  
Best Local Similarity 100.0%; Pred. No. 4.8e-13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 375 EERGVKLGDFIFYSVLVGA 396  
Db 105 EERGVKLGDFIFYSVLVGA 126

RESULT 22  
US-10-051-767-7  
;; Sequence 7, Application US/10051767  
;; Publication No. US20030022151A1  
;; GENERAL INFORMATION:  
;; APPLICANT: THINAKARAN, GOPAL  
;; TITLE OF INVENTION: FUNCTIONAL SCREENING  
;; FILE REFERENCE: ARCD:364US  
;; CURRENT APPLICATION NUMBER: US/10/051,767  
;; CURRENT FILING DATE: 2002-01-17  
;; PRIOR APPLICATION NUMBER: 60/262,353  
;; PRIOR FILING DATE: 2001-01-17  
;; NUMBER OF SEQ ID NOS: 15  
;; SOFTWARE: Patentln Ver. 2.1  
;; SEQ ID NO 7  
;; LENGTH: 166  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
;; OTHER INFORMATION: Peptide  
US-10-051-767-7

Query Match 3.9%; Score 18; DB 9; Length 166;  
Best Local Similarity 100.0%; Pred. No. 3.1e-09;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 375 EERGVKLGDFIFYSVL 392  
Db 142 EERGVKLGDFIFYSVL 159

RESULT 23  
US-10-051-767-13  
;; Sequence 13, Application US/10051767  
;; Publication No. US20030022151A1  
;; GENERAL INFORMATION:  
;; APPLICANT: THINAKARAN, GOPAL  
;; TITLE OF INVENTION: FUNCTIONAL SCREENING  
;; FILE REFERENCE: ARCD:364US  
;; CURRENT APPLICATION NUMBER: US/10/051,767  
;; CURRENT FILING DATE: 2002-01-17  
;; PRIOR APPLICATION NUMBER: 60/262,353  
;; PRIOR FILING DATE: 2001-01-17  
;; NUMBER OF SEQ ID NOS: 15  
;; SOFTWARE: Patentln Ver. 2.1  
;; SEQ ID NO 13  
;; LENGTH: 166  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
;; OTHER INFORMATION: Peptide  
US-10-051-767-13

Query Match 3.9%; Score 18; DB 9; Length 166;

Best Local Similarity 100.0%; Pred. No. 3.1e-09;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 375 EERGKVLGDFIFYSVL 392  
|||||  
Db 142 EERGKVLGDFIFYSVL 159

RESULT 24  
US-10-051-767-12  
Sequence 12, Application US/10051767  
Publication No. US20030022151A1  
GENERAL INFORMATION:  
APPLICANT: THINAKARAN, GOPAL  
TITLE OF INVENTION: FUNCTIONAL SCREENING  
FILE REFERENCE: ARCD:364US  
CURRENT APPLICATION NUMBER: US/10/051,767  
CURRENT FILING DATE: 2002-01-17  
PRIORITY FILING DATE: 2001-01-17  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 12  
LENGTH: 210  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-10-051-767-12

Query Match 3.9%; Score 18; DB 9; Length 210;  
Best Local Similarity 100.0%; Pred. No. 3.9e-09;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 380 KLGLGDFIFYSVLGKAS 397  
|||||  
Db 191 KLGLGDFIFYSVLGKAS 208

RESULT 25  
US-09-823-153-6  
Sequence 6, Application US/09823153  
Patent No. US20020025540A1  
GENERAL INFORMATION:  
APPLICANT: Bristol-Myers Squibb Company  
APPLICANT: Roberts, Susan  
APPLICANT: Pak, Roger  
APPLICANT: Lewis, Martin  
APPLICANT: Smith, David  
APPLICANT: Hendrick, Joseph  
APPLICANT: Vinitsky, Alexander  
TITLE OF INVENTION: ISOLATION OF FUNCTIONALLY ACTIVE GAMMA-SECRETASE PROTEIN COMPLEX  
TITLE OF INVENTION: METHODS FOR DETECTION OF ACTIVITY AND INHIBITORS THEREOF  
FILE REFERENCE: D0004  
CURRENT APPLICATION NUMBER: US/09/823,153  
CURRENT FILING DATE: 2001-07-02  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 6  
LENGTH: 17  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: misc-feature  
OTHER INFORMATION: Description of Artificial Sequence: PSI PEPTIDE  
US-09-823-153-6

Query Match 3.4%; Score 16; DB 10; Length 17;  
Best Local Similarity 100.0%; Pred. No. 2.6e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 45 GHPEPLSNGRPGNSR 60

Db 2 GHPEPLSNGRPGNSR 17  
|||||

RESULT 26  
US-09-823-153-9  
Sequence 9, Application US/09823153  
Patent No. US20020025540A1  
GENERAL INFORMATION:  
APPLICANT: Bristol-Myers Squibb Company  
APPLICANT: Roberts, Susan  
APPLICANT: Pak, Roger  
APPLICANT: Lewis, Martin  
APPLICANT: Smith, David  
APPLICANT: Hendrick, Joseph  
APPLICANT: Vinitsky, Alexander  
TITLE OF INVENTION: ISOLATION OF FUNCTIONALLY ACTIVE GAMMA-SECRETASE PROTEIN COMPL  
TITLE OF INVENTION: METHODS FOR DETECTION OF ACTIVITY AND INHIBITORS THEREOF  
FILE REFERENCE: D0004  
CURRENT APPLICATION NUMBER: US/09/823,153  
CURRENT FILING DATE: 2001-07-02  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 9  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: misc-feature  
OTHER INFORMATION: Description of Artificial Sequence: PSI PEPTIDE  
US-09-823-153-9

Query Match 3.2%; Score 15; DB 10; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.9e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 344 RDSHLGPHRSTPESR 358  
|||||  
Db 1 RDSHLGPHRSTPESR 15

RESULT 27  
US-09-823-153-5  
Sequence 5, Application US/09823153  
Patent No. US20020025540A1  
GENERAL INFORMATION:  
APPLICANT: Bristol-Myers Squibb Company  
APPLICANT: Roberts, Susan  
APPLICANT: Pak, Roger  
APPLICANT: Lewis, Martin  
APPLICANT: Smith, David  
APPLICANT: Hendrick, Joseph  
APPLICANT: Vinitsky, Alexander  
TITLE OF INVENTION: ISOLATION OF FUNCTIONALLY ACTIVE GAMMA-SECRETASE PROTEIN COMPL  
TITLE OF INVENTION: METHODS FOR DETECTION OF ACTIVITY AND INHIBITORS THEREOF  
FILE REFERENCE: D0004  
CURRENT APPLICATION NUMBER: US/09/823,153  
CURRENT FILING DATE: 2001-07-02  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 5  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: misc-feature  
OTHER INFORMATION: Description of Artificial Sequence: PSI PEPTIDE  
US-09-823-153-5

Query Match 3.2%; Score 15; DB 10; Length 16;  
Best Local Similarity 100.0%; Pred. No. 2.1e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 344 RDSHLGPHRSTPEER 358  
|||||  
DB 2 RDSHLGPHRSTPEER 16

## RESULT 28

US-09-795-903A-9  
: Sequence 9, Application US/09795903A  
: Patent No. US20020164760A1  
: GENERAL INFORMATION:  
: APPLICANT: Tang, Jordan J.N.  
: APPLICANT: Lin, Xinli  
: TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods  
: FILE REFERENCE: OMRF 179  
: CURRENT APPLICATION NUMBER: US/09/795,903A  
: CURRENT FILING DATE: 2001-02-28  
: PRIOR APPLICATION NUMBER: 09/604,608  
: PRIOR FILING DATE: 2000-06-27  
: PRIOR APPLICATION NUMBER: 60/168,060  
: PRIOR FILING DATE: 1999-11-30  
: PRIOR APPLICATION NUMBER: 60/177,836  
: PRIOR FILING DATE: 2000-01-25  
: PRIOR APPLICATION NUMBER: 60/178,368  
: PRIOR FILING DATE: 2000-01-27  
: PRIOR APPLICATION NUMBER: 60/210,292  
: PRIOR FILING DATE: 2000-06-08  
: NUMBER OF SEQ ID NOS: 31  
: SOFTWARE: Patentln Ver. 2.1  
: SEQ ID NO 9  
: LENGTH: 8  
: TYPE: PRT  
: ORGANISM: Homo sapiens  
US-09-795-903A-9

Query Match 1.7%; Score 8; DB 9; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 295 LVNMAEGD 302  
|||||

DB 1 LVNMAEGD 8

## RESULT 29

US-09-796-264-9  
: Sequence 9, Application US/09796264  
: Patent No. US20020049303A1  
: GENERAL INFORMATION:  
: APPLICANT: Tang, Jordan J.N.  
: APPLICANT: Lin, Xinli  
: TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods  
: FILE REFERENCE: OMRF 179  
: CURRENT APPLICATION NUMBER: US/09/796,264  
: CURRENT FILING DATE: 2001-02-28  
: PRIOR APPLICATION NUMBER: 09/604,608  
: PRIOR FILING DATE: 2000-06-27  
: PRIOR APPLICATION NUMBER: 60/168,060  
: PRIOR FILING DATE: 1999-11-30  
: PRIOR APPLICATION NUMBER: 60/177,836  
: PRIOR FILING DATE: 2000-01-25  
: PRIOR APPLICATION NUMBER: 60/178,368  
: PRIOR FILING DATE: 2000-01-27  
: PRIOR APPLICATION NUMBER: 60/210,292  
: PRIOR FILING DATE: 2000-06-08  
: NUMBER OF SEQ ID NOS: 31  
: SOFTWARE: Patentln Ver. 2.1  
: SEQ ID NO 9  
: LENGTH: 8  
: TYPE: PRT

: ORGANISM: Homo sapiens  
US-09-796-264-9

Query Match 1.7%; Score 8; DB 10; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 295 LVNMAEGD 302  
|||||

DB 1 LVNMAEGD 8

## RESULT 30

US-09-845-226-9  
: Sequence 9, Application US/09845226  
: Patent No. US20020115600A1  
: GENERAL INFORMATION:  
: APPLICANT: Tang, Jordan J.N.  
: APPLICANT: Hong, Lin  
: APPLICANT: Chosh, Arun K.  
: TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof  
: FILE REFERENCE: OMRF 182  
: CURRENT APPLICATION NUMBER: US/09/845,226  
: CURRENT FILING DATE: 2001-04-30  
: PRIOR APPLICATION NUMBER: 09/603,713  
: PRIOR FILING DATE: 2000-06-27  
: PRIOR APPLICATION NUMBER: 60/168,060  
: PRIOR FILING DATE: 1999-11-30  
: PRIOR APPLICATION NUMBER: 60/177,836  
: PRIOR FILING DATE: 2000-01-25  
: PRIOR APPLICATION NUMBER: 60/178,368  
: PRIOR FILING DATE: 2000-01-27  
: PRIOR APPLICATION NUMBER: 60/210,292  
: PRIOR FILING DATE: 2000-06-08  
: NUMBER OF SEQ ID NOS: 31  
: SOFTWARE: Patentln Ver. 2.1  
: SEQ ID NO 9  
: LENGTH: 8  
: TYPE: PRT  
: ORGANISM: Homo sapiens  
US-09-845-226-9

Query Match 1.7%; Score 8; DB 10; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 295 LVNMAEGD 302  
|||||

DB 1 LVNMAEGD 8

## RESULT 31

US-09-982-172-103  
: Sequence 103, Application US/09982172  
: Patent No. US20020137119A1  
: GENERAL INFORMATION:  
: APPLICANT: Emil Israel Katz  
: TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBIO  
: TITLE OF INVENTION: DIRECTED THEREGAINST, AND METHODS, SYSTEMS AND KITS FOR GENE  
: FILE REFERENCE: 01/22283  
: CURRENT APPLICATION NUMBER: US/09/982,172  
: CURRENT FILING DATE: 2001-10-19  
: NUMBER OF SEQ ID NOS: 253  
: SOFTWARE: Patentln version 3.1  
: SEQ ID NO 103  
: LENGTH: 45  
: TYPE: PRT  
: ORGANISM: Artificial sequence  
: FEATURE:  
: OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-103

Query Match 1.7%; Score 8; DB 10; Length 45;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 420 LTLALLAI 427  
| | | | | | | |  
DB 25 LTLALLAI 32

RESULT 32.  
US-09-982-172-227  
Sequence 227, Application US/09982172  
Patent No. US20020137119A1  
GENERAL INFORMATION:  
APPLICANT: Emil Israel Katz  
TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIE  
TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERAT  
TITLE OF INVENTION: UTILIZING EACH  
FILE REFERENCE: 01/22283  
CURRENT APPLICATION NUMBER: US/09/982.172  
CURRENT FILING DATE: 2001-10-19  
NUMBER OF SEQ ID NOS: 253  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 227  
LENGTH: 45  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-227

Query Match 1.7%; Score 8; DB 10; Length 45;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 420 LTLALLAI 427  
| | | | | | | |  
DB 25 LTLALLAI 32

RESULT 33  
US-10-004-717-7  
Sequence 7, Application US/10004717  
Publication No. US2002019265A1  
GENERAL INFORMATION:  
APPLICANT: ZOGHBI, HUDA Y.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN  
TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,  
FILE REFERENCE: P01899054  
CURRENT APPLICATION NUMBER: US/10/004.717  
CURRENT FILING DATE: 2002-08-16  
PRIOR APPLICATION NUMBER: 09/585,645  
PRIOR FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: 60/176,993  
PRIOR FILING DATE: 2000-01-19  
PRIOR APPLICATION NUMBER: 60/137,060  
PRIOR FILING DATE: 1999-06-01  
NUMBER OF SEQ ID NOS: 69  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 7  
LENGTH: 263  
TYPE: PRT  
ORGANISM: Mus musculus  
US-10-004-717-7

Query Match 1.7%; Score 8; DB 9; Length 263;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 66 DEEDEEL 73  
| | | | | | | |

DB 38 DEEDEEL 45

RESULT 34  
US-10-004-717-31  
Sequence 31, Application US/10004717  
Publication No. US2002019265A1  
GENERAL INFORMATION:  
APPLICANT: ZOGHBI, HUDA Y.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN  
TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,  
FILE REFERENCE: P01899054  
CURRENT APPLICATION NUMBER: US/10/004.717  
CURRENT FILING DATE: 2002-08-16  
PRIOR APPLICATION NUMBER: 09/585,645  
PRIOR FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: 60/176,993  
PRIOR FILING DATE: 2000-01-19  
PRIOR APPLICATION NUMBER: 60/137,060  
PRIOR FILING DATE: 1999-06-01  
NUMBER OF SEQ ID NOS: 69  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 31  
LENGTH: 263  
TYPE: PRT  
ORGANISM: Mus musculus  
US-10-004-717-31

Query Match 1.7%; Score 8; DB 9; Length 263;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 66 DEEDEEL 73  
| | | | | | | |  
DB 38 DEEDEEL 45

RESULT 35  
US-10-004-717-38  
Sequence 38, Application US/10004717  
Publication No. US2002019265A1  
GENERAL INFORMATION:  
APPLICANT: ZOGHBI, HUDA Y.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN  
TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,  
FILE REFERENCE: P01899054  
CURRENT APPLICATION NUMBER: US/10/004.717  
CURRENT FILING DATE: 2002-08-16  
PRIOR APPLICATION NUMBER: 09/585,645  
PRIOR FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: 60/176,993  
PRIOR FILING DATE: 2000-01-19  
PRIOR APPLICATION NUMBER: 60/137,060  
PRIOR FILING DATE: 1999-06-01  
NUMBER OF SEQ ID NOS: 69  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 38  
LENGTH: 263  
TYPE: PRT  
ORGANISM: Mus musculus  
US-10-004-717-38

Query Match 1.7%; Score 8; DB 9; Length 263;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 66 DEEDEEL 73  
| | | | | | | |  
DB 38 DEEDEEL 45

RESULT 36  
US-09-769-097-2  
; Sequence 2, Application US/09769097  
; Patent No. US20020055128A1  
; GENERAL INFORMATION:  
; APPLICANT: Kimberly Anne Brun  
; APPLICANT: Richard James Chenery  
; APPLICANT: Harma Ellens  
; APPLICANT: John Anthony Feld  
; APPLICANT: Lin Yue  
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES  
; FILE REFERENCE: GP-50009-C2  
; CURRENT APPLICATION NUMBER: US/09/769,097  
; CURRENT FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 09/208,809  
; PRIOR FILING DATE: 1998-12-09  
; PRIOR APPLICATION NUMBER: 09/156,800  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: US99/20770  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 1272  
; TYPE: PRT  
; ORGANISM: RATTUS RATTUS  
US-09-769-097-2

Query Match 1.7%; Score 8; DB 10; Length 1272;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 420 LTLILLAI 427  
|||||  
Db 849 LTLILLAI 856

RESULT 37  
US-09-769-097-4  
; Sequence 4, Application US/09769097  
; Patent No. US20020055128A1  
; GENERAL INFORMATION:  
; APPLICANT: Kimberly Anne Brun  
; APPLICANT: Richard James Chenery  
; APPLICANT: Harma Ellens  
; APPLICANT: John Anthony Feld  
; APPLICANT: Lin Yue  
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES  
; FILE REFERENCE: GP-50009-C2  
; CURRENT APPLICATION NUMBER: US/09/769,097  
; CURRENT FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 09/208,809  
; PRIOR FILING DATE: 1998-12-09  
; PRIOR APPLICATION NUMBER: 09/156,800  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: US99/20770  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 1272  
; TYPE: PRT  
; ORGANISM: RATTUS RATTUS  
US-09-769-097-4

Query Match 1.7%; Score 8; DB 10; Length 1272;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 420 LTLILLAI 427  
|||||  
Db 849 LTLILLAI 856

RESULT 38  
US-09-866-866A-8  
; Sequence 8, Application US/09866866A  
; Patent No. US20020102244A1  
; GENERAL INFORMATION:  
; APPLICANT: Schuetz, John  
; APPLICANT: Sorrentino, Brian  
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells  
; FILE REFERENCE: 1340-1-021C1P2  
; CURRENT APPLICATION NUMBER: US/09/866,866A  
; CURRENT FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: 09/584,586  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: PCT/US99/11825  
; PRIOR FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: 60/086,988  
; PRIOR FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 8  
; LENGTH: 1276  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-866-866A-8

Query Match 1.7%; Score 8; DB 10; Length 1276;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 420 LTLILLAI 427  
|||||  
Db 853 LTLILLAI 860

RESULT 39  
US-10-072-621-7  
; Sequence 7, Application US/10072621  
; Patent No. US20020169137A1  
; GENERAL INFORMATION:  
; APPLICANT: Reiner, Peter B.  
; APPLICANT: Connop, Bruce P.  
; APPLICANT: Pollard, Michelle  
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION  
; FILE REFERENCE: 100103.402  
; CURRENT APPLICATION NUMBER: US/10/072,621  
; CURRENT FILING DATE: 2002-02-08  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 1280  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-072-621-7

Query Match 1.7%; Score 8; DB 9; Length 1280;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 420 LTLILLAI 427  
|||||  
Db 857 LTLILLAI 864

RESULT 40  
US-10-044-671-2  
; Sequence 2, Application US/10044671  
; Patent No. US20020177147A1

```
;; GENERAL INFORMATION:
;; APPLICANT: Washington State University Research Foundation
;; APPLICANT: Maaley, Katrina
;; APPLICANT: Bentjen, Steven
;; TITLE OF INVENTION: MDRI VARIANTS AND METHODS FOR THEIR USE
;; FILE REFERENCE: 4630-61733
;; CURRENT APPLICATION NUMBER: US/10/044,671
;; PRIOR FILING DATE: 2002-01-10
;; PRIOR APPLICATION NUMBER: US 60/261,578
;; PRIOR FILING DATE: 2001-01-12
;; PRIOR APPLICATION NUMBER: US 60/314,829
;; NUMBER OF SEQ ID NOS: 10
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 2
;; LENGTH: 1280
;; TYPE: PRT
;; ORGANISM: Canis familiaris
US-10-044-671-2
```

```
Query Match 1.7%; Score 8; DB 9; Length 1280;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 420 LTLILLAI 427
|||||
Db 858 LTLILLAI 865
```

```
RESULT 41
```

```
US-09-866-866A-2
; Sequence 2, Application US/098668666A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Schuetz, John
; APPLICANT: Sorrentino, Brian
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1280
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-866A-2
```

```
Query Match 1.7%; Score 8; DB 10; Length 1280;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 420 LTLILLAI 427
|||||
Db 857 LTLILLAI 864
```

```
RESULT 42
```

```
US-09-866-866A-4
; Sequence 4, Application US/098668666A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
```

```
;; CURRENT FILING DATE: 2001-08-30
;; PRIOR APPLICATION NUMBER: 09/584,586
;; PRIOR FILING DATE: 2000-05-31
;; PRIOR APPLICATION NUMBER: PCT/US99/11825
;; PRIOR FILING DATE: 1999-05-27
;; PRIOR APPLICATION NUMBER: 60/086,988
;; PRIOR FILING DATE: 1998-05-28
;; NUMBER OF SEQ ID NOS: 27
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 4
;; LENGTH: 1280
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-866-866A-4
```

```
Query Match 1.7%; Score 8; DB 10; Length 1280;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 420 LTLILLAI 427
|||||
Db 857 LTLILLAI 864
```

```
RESULT 43
```

```
US-09-795-903A-6
; Sequence 6, Application US/09795903A
; Patent No. US20020164760A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Jordan J.N.
; APPLICANT: Lin, Xianli
; APPLICANT: Koelsch, Gerald
; TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods
; FILE REFERENCE: OMRP 179
; CURRENT APPLICATION NUMBER: US/09/795,903A
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/604,608
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/168,060
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: 60/177,836
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 60/178,368
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 60/210,292
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-795-903A-6
```

```
Query Match 1.5%; Score 7; DB 9; Length 8;
Best Local Similarity 100.0%; Pred. No. 11e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 296 VNNAECD 302
|||||
Db 2 VNNAECD 8
```

```
RESULT 44
```

```
US-09-796-264-6
; Sequence 6, Application US/09796264
; Patent No. US20020049303A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Jordan J.N.
```



APPLICANT: Lin, Xinli  
APPLICANT: Koelisch, Gerald  
TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods  
TITLE OF INVENTION: of Use Thereof  
FILE REFERENCE: OMRF 179  
CURRENT APPLICATION NUMBER: US/09/796,264  
CURRENT FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: 09/604,608  
PRIOR FILING DATE: 2000-06-27  
PRIOR APPLICATION NUMBER: 60/168,060  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: 60/177,836  
PRIOR FILING DATE: 2000-01-25  
PRIOR APPLICATION NUMBER: 60/178,368  
PRIOR FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: 60/210,292  
PRIOR FILING DATE: 2000-06-08  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 8  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-796-264-6

Query Match 1.5%; Score 7; DB 10; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 VNMAGCD 302  
DB 2 VNMAGCD 8

RESULT 45  
US-09-845-226-6  
Sequence 6, Application US/09845226  
Patent No. US20020115600A1  
GENERAL INFORMATION:  
APPLICANT: Tang, Jordan J.N.  
APPLICANT: Hong, Lin  
APPLICANT: Chosh, Arun K.  
TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof  
FILE REFERENCE: OMRF 182  
CURRENT APPLICATION NUMBER: US/09/845,226  
CURRENT FILING DATE: 2001-04-30  
PRIOR APPLICATION NUMBER: 09/603,713  
PRIOR FILING DATE: 2000-06-27  
PRIOR APPLICATION NUMBER: 60/168,060  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: 60/177,836  
PRIOR FILING DATE: 2000-01-25  
PRIOR APPLICATION NUMBER: 60/178,368  
PRIOR FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: 60/210,292  
PRIOR FILING DATE: 2000-06-08  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 8  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-845-226-6

Query Match 1.5%; Score 7; DB 10; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 VNMAGCD 302  
DB 2 VNMAGCD 8

RESULT 46  
US-09-864-761-38169  
Sequence 38169, Application US/09864761  
Patent No. US20020048765A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aecmca-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 38169  
LENGTH: 25  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL050342.36  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.91  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.95  
US-09-864-761-38169

Query Match 1.5%; Score 7; DB 10; Length 25;  
Best Local Similarity 100.0%; Pred. No. 8.1;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 171 LLLPFF 177

DB 9 LLLLLLFF 15

```

RESULT 47
US-09-864-761-38937
; Sequence 38937, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38937
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005848.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.9
; OTHER INFORMATION: EST_HUMAN HIT: BE018257.1, EVALUATE 9.00e-21
; OTHER INFORMATION: SWISSPROT HIT: Q01970, EVALUATE 7.00e-22
US-09-864-761-38937

```

Query Match 1.5%; Score 7; DB 10; Length 55;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 66 DEEDEE 72  
|||||  
DB 48 DEEDEE 54

```

RESULT 48
US-09-764-869-1215
; Sequence 1215, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1215
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (18)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-869-1215

```

Query Match 1.5%; Score 7; DB 10; Length 58;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 168 ISSLLLL 174  
|||||  
DB 22 ISSLLLL 28

```

RESULT 49
US-09-814-122-39
; Sequence 39, Application US/09814122
; Patent No. US20020058307A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen et al.
; TITLE OF INVENTION: 20 Human Secreted Proteins
; FILE REFERENCE: P2005P1
; CURRENT APPLICATION NUMBER: US/09/814,122
; CURRENT FILING DATE: 2001-03-22
; EARLIER APPLICATION NUMBER: US/09/166,780
; EARLIER FILING DATE: 1998-10-06
; EARLIER APPLICATION NUMBER: PCT/US98/06801
; EARLIER FILING DATE: 1998-04-07
; EARLIER APPLICATION NUMBER: 60/042,726
; EARLIER FILING DATE: 1997-04-08
; EARLIER APPLICATION NUMBER: 60/042,727
; EARLIER FILING DATE: 1997-04-08
; EARLIER APPLICATION NUMBER: 60/042,728
; EARLIER FILING DATE: 1997-04-08
; EARLIER APPLICATION NUMBER: 60/042,754
; EARLIER FILING DATE: 1997-04-08
; EARLIER APPLICATION NUMBER: 60/042,825
; EARLIER FILING DATE: 1997-04-08
; EARLIER APPLICATION NUMBER: 60/048,068
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,070
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,184
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 86

```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (51)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-814-122-39

Query Match
Best Local Similarity 1.5%; Score 7; DB 10; Length 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 ISSLLLL 174
DB 5 ISSLLLL 11

RESULT 50
US-09-864-761-47359
; Sequence 47359, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemolca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
```

```
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47359
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005045.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.48
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: SWISSPROT HIT: P21439, EVALUATE 2.00e-27
US-09-864-761-47359

Query Match
Best Local Similarity 1.5%; Score 7; DB 10; Length 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 420 LTLALLA 426
DB 30 LTLALLA 36
```

Search completed: February 6, 2003, 14:27:21  
Job time : 41.4398 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: February 6, 2003, 14:17:31 : Search time 28.0975 Seconds  
(without alignments)  
1597.820 Million cell updates/sec

Title: US-09-689-159a-2

Perfect score: 467

Sequence: 1 MTELPAPLSTYFQNMQSESDN.....ATDYLVQPFMDQLAFHQFYI 467

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	366	78.4	467	2	S58396
2	337	72.2	463	2	S63683
3	188	40.3	374	2	S63683
4	107	22.9	463	2	JC5081
5	107	22.9	467	2	JC5080
6	97	20.8	467	2	I78388
7	59	12.6	433	2	JC5390
8	28	6.0	442	2	I39174
9	28	6.0	448	2	A56993
10	27	5.8	449	2	JC5391
11	18	3.9	461	2	S60253
12	18	3.9	836	2	E89453
13	15	3.2	397	2	A84702
14	15	3.2	453	2	T00724
15	9	1.9	334	2	G90492
16	8	1.7	103	2	H65055
17	8	1.7	103	2	B91079
18	8	1.7	103	2	C85924
19	8	1.7	176	2	A49304
20	8	1.7	314	2	JC7335
21	8	1.7	317	2	AF0205
22	8	1.7	347	2	H75253
23	8	1.7	354	2	T39599
24	8	1.7	392	2	A84564
25	8	1.7	581	2	A97119
26	8	1.7	707	2	A42322
27	8	1.7	1022	2	I53078
28	8	1.7	1104	1	DVMS1A
29	8	1.7	1276	1	DVHY1C

30	1.7	1276	2	A34786	multidrug resistan
31	1.7	1280	1	DVHU1	multidrug resistan
32	1.7	1613	2	S39059	protein Bng1 - hum
33	1.7	1726	2	A39401	merozoite surface
34	1.5	49	2	A56388	dopamine- and camp
35	1.5	65	2	S59962	antimicrobial pept
36	1.5	66	1	A53744	ranalexin precursor
37	1.5	72	1	LMHW3	H+-transporting tw
38	1.5	78	2	C71978	hypothetical prote
39	1.5	85	2	B47354	ygdr protein - sal
40	1.5	86	2	C71549	hypothetical prote
41	1.5	103	2	AF0856	conserved hypothet
42	1.5	104	2	C83236	hypothetical prote
43	1.5	107	2	H72575	hypothetical prote
44	1.5	108	2	D72617	hypothetical prote
45	1.5	109	2	B72763	hypothetical prote
46	1.5	117	2	S31130	hypothetical prote
47	1.5	121	2	S61635	ribosomal protein
48	1.5	125	2	E97486	hypothetical prote
49	1.5	125	2	AE2704	two component resp
50	1.5	131	2	D87491	hypothetical prote
51	1.5	137	2	A97989	hypothetical prote
52	1.5	167	2	B71518	probable signal pe
53	1.5	168	2	T03168	hypothetical prote
54	1.5	169	2	S56281	probable membrane
55	1.5	169	2	H72611	hypothetical prote
56	1.5	170	2	A64316	hypothetical prote
57	1.5	173	1	S22538	oleosin, 18.5K - A
58	1.5	175	2	S70915	major oleosin nap-
59	1.5	176	2	A83847	cytochrome c bioge
60	1.5	183	2	S25089	oleosin Bn-V - rap
61	1.5	184	2	T00431	hypothetical prote
62	1.5	188	2	T29936	hypothetical prote
63	1.5	194	2	F72782	hypothetical prote
64	1.5	208	2	P00278	hypothetical prote
65	1.5	210	2	S62355	high mobility grou
66	1.5	210	2	G83929	hypothetical prote
67	1.5	211	2	H90046	hypothetical prote
68	1.5	213	1	S19158	ubiquitin-protein
69	1.5	218	2	AC1253	glycine betaine/ca
70	1.5	218	2	A11615	membrane protein M
71	1.5	226	2	D49591	membrane glycoprot
72	1.5	226	2	S37434	probable ABC-type
73	1.5	226	2	PC6047	antiquitin - rat (
74	1.5	228	2	B54676	phospholipase C (E
75	1.5	239	2	S68257	hypothetical prote
76	1.5	243	2	S76271	hypothetical prote
77	1.5	243	2	T46461	unknown protein (i
78	1.5	252	2	C96556	hypothetical prote
79	1.5	252	2	B84609	hypothetical prote
80	1.5	259	2	G75186	hypothetical prote
81	1.5	262	2	B83827	hypothetical prote
82	1.5	265	2	T15486	hypothetical prote
83	1.5	271	2	H72261	conserved hypothet
84	1.5	274	2	T25404	hypothetical prote
85	1.5	280	2	A48921	conserved hypothet
86	1.5	286	2	A34599	DNA-binding protei
87	1.5	291	2	E82258	cell division prot
88	1.5	296	2	G64877	peptide transport
89	1.5	296	2	E90862	hypothetical prote
90	1.5	296	2	D85756	hypothetical prote
91	1.5	299	2	D84519	probable endoxylol
92	1.5	317	2	B98208	probable uspa prot
93	1.5	317	2	AG3078	hypothetical prote
94	1.5	318	2	T11260	NADH2 dehydrogenas
95	1.5	323	2	A70029	hypothetical prote
96	1.5	323	2	JC2578	DNA repair protein
97	1.5	326	2	S73304	beta transducin-ii
98	1.5	327	2	F87541	hypothetical prote
99	1.5	330	2	A86355	hypothetical prote
100	1.5	333	2	A99990	RNA 3' phosphate cy
101	1.5	339	2	S04403	geranylgeranyl-dip
102	1.5	340	2	G69474	conserved hypothet

103	7	1.5	341	2	AD3145	176	7	1.5	460	2	D84486	hypothetical prote
104	7	1.5	342	2	S72529	177	7	1.5	460	2	A86180	hypothetical prote
105	7	1.5	343	2	E95911	178	7	1.5	462	1	TVMSMC	transforming prote
106	7	1.5	344	2	C01628	179	7	1.5	462	1	TVMSM2	transforming prote
107	7	1.5	347	2	S47871	180	7	1.5	462	2	S20073	transforming prote
108	7	1.5	348	2	T11429	181	7	1.5	464	1	TVHWM2	transforming prote
109	7	1.5	350	2	S51494	182	7	1.5	465	2	S69038	hypothetical prote
110	7	1.5	351	1	F2KKD2	183	7	1.5	470	2	T34447	hypothetical prote
111	7	1.5	351	2	S78348	184	7	1.5	477	2	S47889	TRK system potassi
112	7	1.5	351	2	S42646	185	7	1.5	479	2	A69036	site-specific DNA-
113	7	1.5	351	2	S73278	186	7	1.5	482	2	S72471	probable tRNA-guan
114	7	1.5	351	2	AC2342	187	7	1.5	482	2	A12259	hypothetical prote
115	7	1.5	352	2	AD2374	188	7	1.5	490	2	A96556	probable ABC trans
116	7	1.5	352	2	T07229	189	7	1.5	493	2	T18789	secreted glycoprot
117	7	1.5	352	2	JT0321	190	7	1.5	497	2	G95398	protein Y51HA.5 [
118	7	1.5	352	2	PS0097	191	7	1.5	498	1	VGREGX	probable sugar upt
119	7	1.5	352	2	T06935	192	7	1.5	498	2	F88359	probable transmem
120	7	1.5	353	1	F2NMD2	193	7	1.5	498	2	B95998	hypothetical prote
121	7	1.5	353	1	F2RMD2	194	7	1.5	512	2	S73457	experantia 2 - fr
122	7	1.5	353	1	F2LMD2	195	7	1.5	514	2	A95274	TRK system potassi
123	7	1.5	353	1	F2SPD2	196	7	1.5	517	2	A95274	site-specific DNA-
124	7	1.5	353	1	F2PMD2	197	7	1.5	522	2	T27101	probable tRNA-guan
125	7	1.5	353	1	F2PMD2	198	7	1.5	522	2	A75412	hypothetical prote
126	7	1.5	353	2	S00929	199	7	1.5	536	1	S66087	conserved hypotet
127	7	1.5	353	2	S03435	200	7	1.5	536	1	A47190	amino acid transpo
128	7	1.5	353	2	S58536	201	7	1.5	537	2	E84318	transducer protein
129	7	1.5	353	2	S52397	202	7	1.5	542	2	T04822	HttI transducer [I
130	7	1.5	353	2	T07549	203	7	1.5	555	2	S72474	hypothetical prote
131	7	1.5	353	2	H69418	204	7	1.5	555	2	H89957	site-specific DNA-
132	7	1.5	353	4	S52396	205	7	1.5	581	2	JDVL64	formyltetrahydrofo
133	7	1.5	355	2	F90391	206	7	1.5	591	2	T16915	DNA-directed DNA p
134	7	1.5	357	2	AD2079	207	7	1.5	606	2	S70358	hypothetical prote
135	7	1.5	360	2	G98142	208	7	1.5	610	2	D84104	hypothetical prote
136	7	1.5	361	2	S25063	209	7	1.5	630	2	S29796	oligopeptide ABC t
137	7	1.5	367	2	A10293	210	7	1.5	632	2	T18692	hypothetical prote
138	7	1.5	368	2	T06093	211	7	1.5	632	2	T27057	hypothetical prote
139	7	1.5	370	1	D22930	212	7	1.5	665	2	S75291	cell division prot
140	7	1.5	375	2	JX0131	213	7	1.5	671	2	F86394	hypothetical prote
141	7	1.5	375	2	F71234	214	7	1.5	684	2	T51789	hypothetical prote
142	7	1.5	376	2	AG1420	215	7	1.5	685	2	T22223	protein T24P13.20
143	7	1.5	376	2	AH1795	216	7	1.5	687	1	FYF6W	sel-1 protein - Ca
144	7	1.5	377	2	S30088	217	7	1.5	687	2	B70515	white protein - fr
145	7	1.5	379	1	Z29750	218	7	1.5	689	2	T52060	hypothetical prote
146	7	1.5	379	1	ZPBPMT	219	7	1.5	694	1	DNCHNL	protein MEDA [imp
147	7	1.5	380	2	E67268	220	7	1.5	713	2	S29795	nucleolin - chicke
148	7	1.5	380	2	S66728	221	7	1.5	721	2	B29928	nucleolin - Chines
149	7	1.5	385	2	E85506	222	7	1.5	726	2	B29928	hypothetical prote
150	7	1.5	392	1	A70406	223	7	1.5	731	2	B82922	phosphate transpor
151	7	1.5	396	2	I50707	224	7	1.5	731	2	T51790	hypothetical prote
152	7	1.5	397	2	T35713	225	7	1.5	740	2	T03847	SNF1-binding protei
153	7	1.5	397	2	S21780	226	7	1.5	746	2	S62365	SNF1-related prote
154	7	1.5	401	2	T48576	227	7	1.5	748	2	F91089	PMS system transcr
155	7	1.5	404	2	S65991	228	7	1.5	748	2	A00866	phosphoenolpyruvat
156	7	1.5	406	2	S59296	229	7	1.5	748	2	F85065	PIS system transcr
157	7	1.5	409	2	C87091	230	7	1.5	748	2	T48719	phosphotransferase
158	7	1.5	409	2	S26021	231	7	1.5	751	2	F97479	hypothetical prote
159	7	1.5	411	2	S58094	232	7	1.5	778	2	AE2697	non-motile and pha
160	7	1.5	415	2	A70679	233	7	1.5	778	2	T28706	two component sens
161	7	1.5	415	2	C69682	234	7	1.5	779	2	AE2697	exoribonuclease 10
162	7	1.5	417	2	T31809	235	7	1.5	790	2	T25095	hypothetical prote
163	7	1.5	421	2	T19315	236	7	1.5	809	2	AE0374	probable permealase
164	7	1.5	427	2	T50738	237	7	1.5	871	2	T28706	hypothetical prote
165	7	1.5	432	2	T33833	238	7	1.5	879	1	JDVLC	DNA-directed DNA p
166	7	1.5	438	2	AF0963	239	7	1.5	883	1	JDVLC2	chloroplast outer
167	7	1.5	440	1	B70326	240	7	1.5	884	1	JDVL7	DNA-directed DNA p
168	7	1.5	440	2	S60423	241	7	1.5	884	1	JDVLM8	DNA-directed DNA p
169	7	1.5	440	2	C82410	242	7	1.5	884	1	JDVLM8	DNA-directed DNA p
170	7	1.5	444	2	T20600	243	7	1.5	904	2	C83030	exoribonuclease RN
171	7	1.5	449	2	D87284	244	7	1.5	904	2	S32607	titfunctional enzy
172	7	1.5	452	2	H84772	245	7	1.5	906	1	F69552	leucyl-tRNA synthet
173	7	1.5	454	2	S11511	246	7	1.5	932	2	C81730	metallopeptidase,
174	7	1.5	456	1	TVHDMC	247	7	1.5	939	2	D87326	conserved hypotet
175	7	1.5	460	2	S11558	248	7	1.5	964	2	D87326	

249	7	1.5	1051	2	S27002	322	6	1.3	83	2	JN0400	hypothetical 8.8K
250	7	1.5	1085	2	S62516	323	6	1.3	83	2	B83556	hypothetical prote
251	7	1.5	1130	2	T21134	324	6	1.3	83	2	AH3458	uasee (EC 3.5.1.5
252	7	1.5	1160	2	T00272	325	6	1.3	85	2	T32664	hypothetical prote
253	7	1.5	1217	2	A43493	326	6	1.3	85	2	T17882	hypothetical prote
254	7	1.5	1234	2	S52099	327	6	1.3	86	2	S16831	Ig kappa chain V r
255	7	1.5	1237	2	I38994	328	6	1.3	86	2	D98327	cell division topo
256	7	1.5	1245	2	AE1915	329	6	1.3	86	2	A12955	cell division topo
257	7	1.5	1254	2	D71613	330	6	1.3	87	2	F45714	probable regulator
258	7	1.5	1255	2	E68160	331	6	1.3	87	2	D61547	hypothetical prote
259	7	1.5	1262	2	T33074	332	6	1.3	87	2	H96010	probable cell divi
260	7	1.5	1273	1	VDRLT3	333	6	1.3	88	2	S66287	sapacin B precursor
261	7	1.5	1279	1	DHRTU3	334	6	1.3	88	2	F69528	hypothetical prote
262	7	1.5	1291	1	A28334	335	6	1.3	89	2	B69094	conserved hypotet
263	7	1.5	1298	2	I54367	336	6	1.3	89	2	C86497	hypothetical prote
264	7	1.5	1302	2	A41249	337	6	1.3	89	2	B86833	hypothetical prote
265	7	1.5	1304	1	A46546	338	6	1.3	90	2	S74568	hypothetical prote
266	7	1.5	1317	2	B83346	339	6	1.3	90	2	AD3625	cell division inh1
267	7	1.5	1407	1	T00558	340	6	1.3	91	2	C87034	hypothetical prote
268	7	1.5	1410	2	T15661	341	6	1.3	91	2	D85927	partial probable t
269	7	1.5	1444	1	A30588	342	6	1.3	93	1	E64684	ribosomal protein
270	7	1.5	1568	2	T41013	343	6	1.3	96	2	T18006	hypothetical prote
271	7	1.5	1580	2	T26204	344	6	1.3	97	2	S05334	hypothetical prote
272	7	1.5	1641	2	I38614	345	6	1.3	98	2	S75343	hypothetical prote
273	7	1.5	1664	2	F84485	346	6	1.3	98	2	AE1225	probable cobalt tr
274	7	1.5	1761	2	T14289	347	6	1.3	98	2	AG1578	probable cobalt tr
275	7	1.5	2094	2	S33124	348	6	1.3	99	2	T37093	hypothetical prote
276	7	1.5	2157	2	S71461	349	6	1.3	100	2	S42602	uasee (EC 3.5.1.5
277	7	1.5	2175	1	S03170	350	6	1.3	100	2	T50708	uasee (EC 3.5.1.5
278	7	1.5	2276	2	T00076	351	6	1.3	100	2	AE2872	uasee gamma subu
279	7	1.5	2380	2	E71604	352	6	1.3	100	2	H69534	conserved hypotet
280	7	1.5	2448	2	S53999	353	6	1.3	100	2	E43550	hypothetical prote
281	7	1.5	2647	2	A37098	354	6	1.3	101	2	S60434	probable membrane
282	7	1.5	3530	2	A39266	355	6	1.3	101	2	B97851	hypothetical prote
283	7	1.5	4128	2	JC6306	356	6	1.3	102	1	XKPO	metallicarboxypept
284	6	1.3	23	2	S31210	357	6	1.3	102	2	A32264	parathyrosin - hum
285	6	1.3	23	2	I55406	358	6	1.3	102	2	A24846	apolipoprotein A-I
286	6	1.3	29	2	B34490	359	6	1.3	102	2	S20440	bilirubin UDP-gluc
287	6	1.3	29	2	B61613	360	6	1.3	103	2	F70976	hypothetical prote
288	6	1.3	31	2	D34490	361	6	1.3	103	2	S72973	hypothetical prote
289	6	1.3	43	2	S73245	362	6	1.3	103	2	T47718	hypothetical prote
290	6	1.3	46	2	G85759	363	6	1.3	103	2	G82854	hypothetical prote
291	6	1.3	47	2	D87675	364	6	1.3	104	2	D70879	hypothetical prote
292	6	1.3	48	2	D64614	365	6	1.3	105	2	H72125	hypothetical prote
293	6	1.3	48	2	E71901	366	6	1.3	106	2	AD0408	conserved hypotet
294	6	1.3	53	2	H69359	367	6	1.3	106	2	D69447	hypothetical prote
295	6	1.3	53	2	D89989	368	6	1.3	107	2	F97648	uasee gamma chain
296	6	1.3	57	2	AC2265	369	6	1.3	108	2	S69294	hypothetical prote
297	6	1.3	57	2	A82905	370	6	1.3	108	2	JC6041	agc protein precu
298	6	1.3	57	2	AC0191	371	6	1.3	108	2	AB0636	conserved hypotet
299	6	1.3	60	2	C21774	372	6	1.3	108	2	H83303	hypothetical prote
300	6	1.3	61	2	E39741	373	6	1.3	108	2	G72635	hypothetical prote
301	6	1.3	63	2	T07349	374	6	1.3	109	2	B61030	nitrogen regulator
302	6	1.3	65	2	A75172	375	6	1.3	109	2	D36885	bo-typo ubiquinol
303	6	1.3	65	2	C91082	376	6	1.3	109	2	S69307	probable membrane
304	6	1.3	65	2	S19568	377	6	1.3	109	2	C64402	hypothetical prote
305	6	1.3	67	2	T07256	378	6	1.3	109	2	C82786	conserved hypotet
306	6	1.3	69	2	A87339	379	6	1.3	110	2	F75153	csng protein precu
307	6	1.3	72	2	D97178	380	6	1.3	110	2	S70789	hypothetical prote
308	6	1.3	72	2	AH0039	381	6	1.3	110	2	D68896	hypothetical prote
309	6	1.3	74	2	S54122	382	6	1.3	111	2	S65365	profilin II - bov1
310	6	1.3	75	2	T17767	383	6	1.3	111	2	G70068	hypothetical prote
311	6	1.3	76	2	T17993	384	6	1.3	111	2	T49569	hypothetical prote
312	6	1.3	76	2	A81858	385	6	1.3	112	2	T12731	hypothetical prote
313	6	1.3	77	1	XKTO	386	6	1.3	112	2	AG2539	hypothetical prote
314	6	1.3	77	1	C58652	387	6	1.3	112	2	A72695	high mobility grou
315	6	1.3	77	2	B64390	388	6	1.3	113	2	A43436	hypothetical prote
316	6	1.3	77	2	AH1279	389	6	1.3	113	2	G83978	probable lipid tra
317	6	1.3	78	2	S03763	390	6	1.3	114	1	S20862	lipid transfer pro
318	6	1.3	80	2	T11069	391	6	1.3	114	2	S29227	hypothetical prote
319	6	1.3	80	2	S34219	392	6	1.3	114	2	S32479	conserved hypotet
320	6	1.3	81	2	T47289	393	6	1.3	114	2	A64609	hypothetical prote
321	6	1.3	83	2	S15142	394	6	1.3	115	2	AH1895	hypothetical prote

395	6	1.3	115	2	H72745	hypothetical prote
396	6	1.3	116	2	A39746	NADH2 dehydrogenas
397	6	1.3	116	2	F72493	hypothetical prote
398	6	1.3	118	2	E86282	protein F10B6.28 (
399	6	1.3	119	2	H71042	probable DNA-dirc
400	6	1.3	119	2	T43514	beta-tubulin foldi
401	6	1.3	120	2	S21565	hypothetical prote
402	6	1.3	120	2	G70895	probable PE protei
403	6	1.3	121	2	E85807	unknown protein en
404	6	1.3	121	2	C90959	probable tail prot
405	6	1.3	122	2	S40314	Ig kappa chain - h
406	6	1.3	122	2	S72770	BI496_F1_2 protei
407	6	1.3	122	2	C73571	conserved hypothet
408	6	1.3	122	2	B70397	hypothetical prote
409	6	1.3	124	2	H71230	hypothetical prote
410	6	1.3	125	2	A71024	hypothetical prote
411	6	1.3	127	1	S19498	thioredoxin homolo
412	6	1.3	127	1	F57987	cytochrome c-type
413	6	1.3	127	2	A96261	cytochrome c-type
414	6	1.3	127	2	E86101	hypothetical prote
415	6	1.3	128	2	H81275	hypothetical prote
416	6	1.3	129	2	F84707	hypothetical prote
417	6	1.3	129	2	H86156	hypothetical prote
418	6	1.3	130	1	R35808	ribosomal protein
419	6	1.3	130	1	A69334	mercuric resistanc
420	6	1.3	130	1	A69894	hypothetical prote
421	6	1.3	130	2	C91150	30S ribosomal subu
422	6	1.3	130	2	G85995	30S ribosomal subu
423	6	1.3	130	2	AG1007	30S ribosomal chat
424	6	1.3	130	2	T44809	hypothetical prote
425	6	1.3	130	2	T47291	hypothetical prote
426	6	1.3	131	2	A44990	NADH2 dehydrogenas
427	6	1.3	131	2	PQ0059	T-cell receptor be
428	6	1.3	131	2	E97737	hypothetical prote
429	6	1.3	131	2	T49331	hypothetical prote
430	6	1.3	131	2	AD0744	probable membrane
431	6	1.3	131	2	B96782	unknown protein F2
432	6	1.3	132	2	S17183	interferon-induced
433	6	1.3	132	2	AB1752	hypothetical prote
434	6	1.3	132	2	AH2171	hypothetical prote
435	6	1.3	133	2	G71028	hypothetical prote
436	6	1.3	134	2	I48639	neurotoxin homolog
437	6	1.3	134	2	G82093	hypothetical prote
438	6	1.3	135	2	D71659	hypothetical prote
439	6	1.3	135	2	S74746	hypothetical prote
440	6	1.3	135	2	G72714	hypothetical prote
441	6	1.3	136	2	S36304	T-cell receptor de
442	6	1.3	136	2	S36320	T-cell receptor de
443	6	1.3	136	2	G95224	hypothetical prote
444	6	1.3	136	2	AB3058	hypothetical prote
445	6	1.3	136	2	T08787	probable testin DK
446	6	1.3	137	2	S36303	T-cell receptor de
447	6	1.3	137	2	C71118	hypothetical prote
448	6	1.3	138	2	A33334	Ig heavy chain pre
449	6	1.3	138	2	T07203	hypothetical prote
450	6	1.3	138	2	F72621	hypothetical prote
451	6	1.3	138	2	D95289	hypothetical prote
452	6	1.3	139	2	A38612	insulin-like growt
453	6	1.3	139	2	B75307	hypothetical prote
454	6	1.3	139	2	G82433	conserved hypothet
455	6	1.3	139	2	T26722	hypothetical prote
456	6	1.3	140	2	S36804	profilin II - huma
457	6	1.3	140	2	AE3007	conserved hypothet
458	6	1.3	140	2	G84608	En/Spm-like transp
459	6	1.3	141	2	S36318	T-cell receptor de
460	6	1.3	141	2	A86875	50S ribosomal prot
461	6	1.3	141	2	S54700	PilE protein - pse
462	6	1.3	141	2	E55855	lysA protein - pha
463	6	1.3	141	2	E64949	probable membrane
464	6	1.3	141	2	C90951	hypothetical prote
465	6	1.3	141	2	G85799	hypothetical prote
466	6	1.3	142	2	G82832	aspartate 1-decarb
467	6	1.3	142	2	E72571	hypothetical prote
468	6	1.3	142	2	JC7636	cystatin 1 - wheat
469	6	1.3	142	2	E84396	hypothetical prote
470	6	1.3	142	2	T47562	hypothetical prote
471	6	1.3	142	2	B72667	hypothetical prote
472	6	1.3	143	2	S36321	T-cell receptor de
473	6	1.3	143	2	T22164	hypothetical prote
474	6	1.3	143	2	H90037	hypothetical prote
475	6	1.3	144	1	PSHUYF	phospholipase A2 (
476	6	1.3	144	2	T35720	hypothetical prote
477	6	1.3	144	2	T33736	hypothetical prote
478	6	1.3	144	2	AB3266	asparagine-binding
479	6	1.3	145	1	CCRR25	cytochrome c2 prec
480	6	1.3	145	2	A41047	exoenzyme S synthe
481	6	1.3	145	2	D71868	hypothetical prote
482	6	1.3	145	2	D82105	conserved hypothet
483	6	1.3	146	1	UDHU	cystatin C precurs
484	6	1.3	146	2	G70183	chemotaxis respons
485	6	1.3	146	2	S61393	hypothetical prote
486	6	1.3	146	2	F83814	Na+/H+ antiporter
487	6	1.3	147	2	AD0953	MioC protein (limp
488	6	1.3	147	2	D81404	hypothetical prote
489	6	1.3	147	2	B98334	hypothetical prote
490	6	1.3	147	2	AH2948	hypothetical prote
491	6	1.3	148	1	G69952	conserved hypothet
492	6	1.3	148	2	A72785	hypothetical prote
493	6	1.3	148	2	E87609	hypothetical prote
494	6	1.3	148	2	E90359	partial transposas
495	6	1.3	149	2	A11824	hypothetical prote
496	6	1.3	150	2	S44505	hypothetical 15.9K
497	6	1.3	150	2	C86224	hypothetical prote
498	6	1.3	151	2	A64503	conserved hypothet
499	6	1.3	151	2	A81863	hypothetical prote
500	6	1.3	151	2	G90036	hypothetical prote
501	6	1.3	152	2	S73603	hypothetical prote
502	6	1.3	152	2	H83813	hypothetical prote
503	6	1.3	152	2	C97371	hypothetical prote
504	6	1.3	152	2	AC2589	conserved hypothet
505	6	1.3	153	2	G72302	hypothetical prote
506	6	1.3	153	2	AE2344	hypothetical prote
507	6	1.3	153	2	G81236	smg protein, proba
508	6	1.3	154	2	AG1021	probable acetylta
509	6	1.3	154	2	E82112	conserved hypothet
510	6	1.3	156	2	S41029	hypothetical prote
511	6	1.3	156	2	T28997	hypothetical prote
512	6	1.3	156	2	B64036	hypothetical prote
513	6	1.3	157	2	G02355	tumor-associated m
514	6	1.3	157	2	A10583	conserved hypothet
515	6	1.3	157	2	B47130	high mobility grou
516	6	1.3	157	2	G82273	conserved hypothet
517	6	1.3	158	2	F70918	probable regulator
518	6	1.3	159	2	AG2207	hypothetical prote
519	6	1.3	159	2	F69428	hypothetical prote
520	6	1.3	160	2	T11139	NADH2 dehydrogenas
521	6	1.3	160	2	A96277	hypothetical prote
522	6	1.3	160	2	AC2799	conserved hypothet
523	6	1.3	161	2	G83727	phosphoribosylamin
524	6	1.3	161	2	G71407	transcription fact
525	6	1.3	161	2	S18991	high mobility grou
526	6	1.3	162	2	G70413	NADH2 dehydrogenas
527	6	1.3	162	2	C71000	hypothetical prote
528	6	1.3	163	2	G90503	conserved hypothet
529	6	1.3	164	2	I40187	nimD protein - Bac
530	6	1.3	166	2	A48465	ribosomal protein
531	6	1.3	167	2	G87610	hypothetical prote
532	6	1.3	168	2	E75257	molymphenum cotact
533	6	1.3	168	2	A96654	hypothetical prote
534	6	1.3	168	2	T03640	high mobility grou
535	6	1.3	168	2	C84014	intracellular prot
536	6	1.3	169	2	T49618	2',3'-cyclic-nucle
537	6	1.3	170	2	C70102	hypothetical prote
538	6	1.3	172	2	E72472	hypothetical prote
539	6	1.3	173	2	A48300	lens membrane prot
540	6	1.3	173	2	A49182	integral membrane



541	6	1.3	173	2	G86905	hypothetical prote
542	6	1.3	174	1	RDSPYA	ferredoxin-thione
543	6	1.3	174	2	G83257	general secretion
544	6	1.3	174	2	S12356	xcpz protein - pse
545	6	1.3	174	2	A27293	probable y chromos
546	6	1.3	175	2	JC5365	TBP-binding repres
547	6	1.3	175	2	AD1327	cell-division init
548	6	1.3	175	2	AF1698	platelet glycoprot
549	6	1.3	176	1	A46606	cystatin-related p
550	6	1.3	176	2	UT0616	hypothetical prote
551	6	1.3	178	2	A64399	transcription regu
552	6	1.3	179	2	T40442	hypothetical highl
553	6	1.3	181	2	B75430	hypothetical prote
554	6	1.3	181	2	D81450	aminoacyl-tRNA hyd
555	6	1.3	181	2	T22884	hypothetical prote
556	6	1.3	182	2	H95055	conserved hypotnet
557	6	1.3	182	2	E97925	conserved hypotnet
558	6	1.3	183	2	C85439	probable cytoskele
559	6	1.3	183	2	S76453	hypothetical prote
560	6	1.3	183	2	G96018	probable methylate
561	6	1.3	184	2	S43778	phycoerythrin I be
562	6	1.3	184	2	H45045	phycoerythrin I be
563	6	1.3	184	2	E82124	cytochrome c bioge
564	6	1.3	184	2	T36644	probable transcrip
565	6	1.3	184	2	T08200	probable disease r
566	6	1.3	185	2	T08198	hypothetical prote
567	6	1.3	185	2	H75216	conserved hypotnet
568	6	1.3	185	2	F89961	conserved hypotnet
569	6	1.3	186	2	F81689	probable integrat
570	6	1.3	187	2	F81434	varion protein j13
571	6	1.3	189	2	S52907	hypothetical prote
572	6	1.3	189	2	E84152	probable transcrip
573	6	1.3	189	2	A83586	hypothetical prote
574	6	1.3	191	2	T19026	hypothetical prote
575	6	1.3	191	2	T13536	hypothetical prote
576	6	1.3	191	2	F72277	hypothetical prote
577	6	1.3	192	2	B85040	hypothetical prote
578	6	1.3	193	1	G69347	conserved hypotnet
579	6	1.3	193	2	A10420	phnM protein [limp
580	6	1.3	193	2	A97068	enzyme from phosph
581	6	1.3	193	2	G88039	membrane protein P
582	6	1.3	194	2	S28850	membrane protein P
583	6	1.3	194	2	E71064	hypothetical prote
584	6	1.3	194	2	E75083	hypothetical prote
585	6	1.3	195	2	S50456	pituitary adenylat
586	6	1.3	196	2	D95546	hypothetical prote
587	6	1.3	197	2	T04782	hypothetical prote
588	6	1.3	197	2	F70622	probable transcrip
589	6	1.3	199	2	T39861	m103 protein - fis
590	6	1.3	200	2	G83204	probable peroxidase
591	6	1.3	200	2	T30472	hypothetical prote
592	6	1.3	201	2	C64451	CDPdiacylglycerol-
593	6	1.3	201	2	S10293	HMG-1 - chicken
594	6	1.3	201	2	G64892	[acyl-carrier-prot
595	6	1.3	201	2	F90880	acyl carrier prote
596	6	1.3	201	2	C85738	acyl carrier prote
597	6	1.3	201	2	AH0664	P-factor mating ph
598	6	1.3	201	2	A36985	lipoprotein signal
599	6	1.3	201	2	A87059	probable membrane
600	6	1.3	202	2	S22359	nonhistone chromos
601	6	1.3	202	2	C97472	probable membrane
602	6	1.3	202	2	AG2690	lysE family transp
603	6	1.3	202	2	A26301	dopamine- and camp
604	6	1.3	203	2	C64038	hypothetical prote
605	6	1.3	203	2	G84350	hypothetical prote
606	6	1.3	203	2	S48955	hypothetical prote
607	6	1.3	203	2	S67607	probable membrane
608	6	1.3	204	2	A89199	conserved hypotnet
609	6	1.3	205	2	D84527	probable ADP-ribos
610	6	1.3	205	2	S26062	nonhistone chromos
611	6	1.3	205	2	G71891	hypothetical prote
612	6	1.3	205	2	E64621	conserved hypotnet
613	6	1.3	205	2	F87151	conserved hypotnet
614	6	1.3	205	2	T37103	hypothetical prote
615	6	1.3	205	2	T27923	hypothetical prote
616	6	1.3	205	2	S41002	hypothetical prote
617	6	1.3	205	2	C96998	phosphatidylserine
618	6	1.3	206	2	T31585	hypothetical prote
619	6	1.3	206	2	T41232	hypothetical prote
620	6	1.3	206	2	H84165	hypothetical prote
621	6	1.3	207	2	JC1129	nonhistone chromos
622	6	1.3	207	2	JC1114	high-mobility grou
623	6	1.3	207	2	S74328	hypothetical prote
624	6	1.3	207	2	A82270	acetyltransferase-
625	6	1.3	207	2	B71446	hypothetical prote
626	6	1.3	208	2	S47687	hypothetical prote
627	6	1.3	208	2	E91168	hypothetical prote
628	6	1.3	208	2	E86014	probable enzyme [i
629	6	1.3	208	2	T27398	probable enzyme yh
630	6	1.3	208	2	E86221	hypothetical prote
631	6	1.3	209	2	H75626	DNA-binding respon
632	6	1.3	209	2	B83921	hypothetical prote
633	6	1.3	210	1	S74656	hypothetical prote
634	6	1.3	210	2	S54774	high mobility grou
635	6	1.3	210	2	D44805	hypothetical prote
636	6	1.3	210	2	C84278	hypothetical prote
637	6	1.3	210	2	C81216	conserved hypotnet
638	6	1.3	210	2	G81793	hypothetical prote
639	6	1.3	210	2	G72507	hypothetical prote
640	6	1.3	211	2	C64751	hypothetical prote
641	6	1.3	211	2	E70253	hypothetical prote
642	6	1.3	211	2	T29426	hypothetical prote
643	6	1.3	213	2	D81690	probable sodium-tr
644	6	1.3	213	2	A88042	protein F56D12.2 [
645	6	1.3	213	2	T30116	hypothetical prote
646	6	1.3	214	2	T47110	3-oxoadipate CoA-t
647	6	1.3	214	2	S76538	hypothetical prote
648	6	1.3	215	1	UOXFAS	ubiquitin-protein
649	6	1.3	215	1	S01947	nonhistone chromos
650	6	1.3	215	2	A28897	nonhistone chromos
651	6	1.3	215	2	S02826	nonhistone chromos
652	6	1.3	215	2	B97858	mat protein (limpor
653	6	1.3	215	2	H64924	hypothetical prote
654	6	1.3	215	2	C90926	hypothetical prote
655	6	1.3	215	2	G85774	hypothetical prote
656	6	1.3	216	2	S29857	nonhistone chromos
657	6	1.3	216	2	S21337	genome polyprotein
658	6	1.3	216	2	G70474	hypothetical prote
659	6	1.3	217	2	T35020	hypothetical prote
660	6	1.3	217	2	T14621	butyrate-acetoacet
661	6	1.3	217	2	S50818	FE65 nuclear prote
662	6	1.3	218	2	D71693	hypothetical prote
663	6	1.3	218	2	E89004	protein T24A6.12 [
664	6	1.3	218	2	E70818	probable mercury t
665	6	1.3	218	2	AG3554	multiple sugar-bin
666	6	1.3	219	2	T10509	phosphoglycolate p
667	6	1.3	219	2	C86229	hypothetical prote
668	6	1.3	219	2	F71363	probable amino aci
669	6	1.3	219	2	A10242	conserved hypotnet
670	6	1.3	219	2	A13200	dehydratogenase Atus
671	6	1.3	220	1	TEBE12	legument protein -
672	6	1.3	220	2	T42619	FLT3/FLK2 ligand (
673	6	1.3	220	2	S43291	flt3 ligand isoform
674	6	1.3	220	2	S58343	competence protein
675	6	1.3	220	2	A95258	hypothetical prote
676	6	1.3	220	2	T14951	hypothetical prote
677	6	1.3	220	2	A99523	hypothetical prote
678	6	1.3	221	2	AF3483	heme exporter prot
679	6	1.3	221	2	H72466	hypothetical prote
680	6	1.3	221	2	D75179	hypothetical prote
681	6	1.3	221	2	T23821	hypothetical prote
682	6	1.3	221	1	D86878	hypothetical prote
683	6	1.3	222	1	S25244	regulatory protein
684	6	1.3	222	2	H71642	mat protein (mat)
685	6	1.3	222	2	D82132	hypothetical prote
686	6	1.3	223	1	B48640	regulatory protein

687	6	1.3	223	2	S60088	nitrate reductase	760	6	1.3	246	2	C86033	probable outer mem
688	6	1.3	223	2	AH2354	nitrogen-responsiv	761	6	1.3	246	2	AB1753	Orf53 [bacterioph
689	6	1.3	223	2	T05755	hypothetical prote	762	6	1.3	247	2	B64873	probable membrane
690	6	1.3	224	2	S71749	DCL protein precur	763	6	1.3	247	2	C90848	hypothetical prote
691	6	1.3	225	1	A48640	global nitrogen re	764	6	1.3	247	2	B85706	hypothetical prote
692	6	1.3	225	2	T01175	hypothetical prote	765	6	1.3	247	2	T39909	hypothetical prote
693	6	1.3	225	2	B70150	hypothetical prote	766	6	1.3	247	2	I59334	leucine-rich acid
694	6	1.3	226	2	T35202	probable two-compo	767	6	1.3	247	2	E90127	hypothetical prote
695	6	1.3	226	2	D72769	ATP synthase F0 ch	768	6	1.3	247	2	AB2465	hypothetical prote
696	6	1.3	227	2	P90622	ATP synthase F0 ch	769	6	1.3	248	1	G64839	ymcB protein - Esc
697	6	1.3	227	2	S04133	photosystem I chal	770	6	1.3	248	2	E90771	hypothetical prote
698	6	1.3	228	1	QXASC	NADH dehydrogenas	771	6	1.3	248	2	A85634	hypothetical prote
699	6	1.3	228	1	C64484	very hypothetical	772	6	1.3	248	2	AH3575	surface protein [l
700	6	1.3	230	2	T37811	hypothetical prote	773	6	1.3	249	1	A35842	chymase (EC 3.4.21
701	6	1.3	231	2	D90966	hypothetical prote	774	6	1.3	249	2	S43309	probable HLA class
702	6	1.3	231	2	D85814	hypothetical prote	775	6	1.3	249	2	C75156	hypothetical prote
703	6	1.3	231	2	A49265	flk/fik-2 ligand	776	6	1.3	249	2	F91274	hypothetical prote
704	6	1.3	232	2	A86720	conserved hypotet	777	6	1.3	249	2	F86115	hypothetical prote
705	6	1.3	232	2	E87681	conserved hypotet	778	6	1.3	249	2	S56415	hypothetical prote
706	6	1.3	232	2	B75561	conserved hypotet	779	6	1.3	250	2	AB1722	ABC transporter (A
707	6	1.3	232	2	D81274	N-acylneuraminate	780	6	1.3	250	2	AG1351	ABC transporter (A
708	6	1.3	232	2	A27503	testis-specific pr	781	6	1.3	250	2	C90449	conserved hypotet
709	6	1.3	232	2	E97347	probable transcrip	782	6	1.3	251	1	JH0147	uracil phosphoribo
710	6	1.3	233	2	AF0171	conserved hypotet	783	6	1.3	251	2	A91067	transposase of ins
711	6	1.3	233	2	T26133	hypothetical prote	784	6	1.3	251	2	S75353	hypothetical prote
712	6	1.3	234	2	S35020	nolW protein - Rhl	785	6	1.3	251	2	A85911	IS30 transposase (
713	6	1.3	236	2	A28566	T-cell suppressor	786	6	1.3	252	2	T14964	phage lambda-relat
714	6	1.3	236	2	G75311	branched-chain aml	787	6	1.3	252	2	E97070	cobalamin-5-phosph
715	6	1.3	237	2	T22859	hypothetical prote	788	6	1.3	253	2	BVECCZ	probable sulfate t
716	6	1.3	237	2	B70529	hypothetical prote	789	6	1.3	253	2	B71547	probable typtopha
717	6	1.3	237	2	T15389	hypothetical prote	790	6	1.3	253	2	S07264	beta-crystallin B1
718	6	1.3	237	2	E75169	hypothetical prote	791	6	1.3	253	2	E91039	required for sulfa
719	6	1.3	238	2	H64016	hypothetical prote	792	6	1.3	253	2	H85883	conserved hypotet
720	6	1.3	238	2	S48984	hypothetical prote	793	6	1.3	253	2	C83148	hypothetical prote
721	6	1.3	238	2	G87577	Mutr/nudix family	794	6	1.3	253	2	H86725	hypothetical prote
722	6	1.3	238	2	D96026	probable branched-	795	6	1.3	254	2	T10185	MADS-box protein C
723	6	1.3	238	2	H89630	hypothetical prote	796	6	1.3	254	2	H69057	hypothetical prote
724	6	1.3	238	2	AC3588	high-affinity bran	797	6	1.3	255	2	A95181	dihydrodipicolinat
725	6	1.3	238	2	AD2138	hypothetical prote	798	6	1.3	255	2	E98048	cytochrome-c oxida
726	6	1.3	239	2	S52823	hypothetical prote	799	6	1.3	255	2	S62759	hypothetical prote
727	6	1.3	239	2	T10075	pepsin inhibitor-1	800	6	1.3	255	2	T21686	hypothetical prote
728	6	1.3	239	2	B83709	hypothetical prote	801	6	1.3	255	2	H71966	flagellar biosynth
729	6	1.3	239	2	AC2657	conserved hypotet	802	6	1.3	255	2	E64541	flagellar biosynth
730	6	1.3	240	2	A97531	branched-chain aml	803	6	1.3	255	2	B82107	probable flagellar
731	6	1.3	240	2	AB2750	hepatoma-derived g	804	6	1.3	256	2	C72261	spermidine/putresc
732	6	1.3	240	2	A55055	probable regulator	805	6	1.3	256	2	A82595	NADH pyrophosphat
733	6	1.3	240	2	E45714	hypothetical prote	806	6	1.3	257	1	G65206	hypothetical prote
734	6	1.3	240	2	T46180	SNF7 protein - yea	807	6	1.3	257	2	D96016	hypothetical prote
735	6	1.3	240	2	S52590	probable transfera	808	6	1.3	257	2	G91243	conserved hypotet
736	6	1.3	241	2	H70561	hypothetical prote	809	6	1.3	257	2	E86091	hypothetical prote
737	6	1.3	241	2	G71045	hypothetical prote	810	6	1.3	257	2	AD0932	probable flagellar
738	6	1.3	241	2	C64006	probable S-adenosy	811	6	1.3	257	2	AC0087	carboxyphosphoen
739	6	1.3	241	2	C46181	PX-tax-orf II (alt	812	6	1.3	257	2	AD1441	pancreatic elastas
740	6	1.3	241	2	AB1131	hypothetical prote	813	6	1.3	258	2	S70439	hypothetical prote
741	6	1.3	241	2	AB1491	hypothetical prote	814	6	1.3	258	2	D83779	hypothetical prote
742	6	1.3	241	2	B75628	DNA-binding stress	815	6	1.3	258	2	G71423	hypothetical prote
743	6	1.3	241	2	G69899	transcription regu	816	6	1.3	259	1	PMHUBM	bisphosphoglycerat
744	6	1.3	241	2	T09769	uridylate kinase (	817	6	1.3	259	1	PMRBBM	bisphosphoglycerat
745	6	1.3	243	2	E81080	conserved hypotet	818	6	1.3	259	1	PMMSBM	bisphosphoglycerat
746	6	1.3	243	2	B97423	transcription acti	819	6	1.3	259	2	C95038	PTS system, ITC co
747	6	1.3	243	2	AH2640	transcription regu	820	6	1.3	259	2	F97908	hypothetical prote
748	6	1.3	243	2	AI3227	geneC' protein [lm	821	6	1.3	259	2	T43102	hypothetical prote
749	6	1.3	244	2	F64247	triose-phosphate 1	822	6	1.3	259	2	A71537	probable n-acetylm
750	6	1.3	244	2	S73539	triosephosphate 1s	823	6	1.3	260	2	G84714	hypothetical prote
751	6	1.3	244	2	C96796	probable heat choc	824	6	1.3	261	2	G58892	cytochrome-c oxida
752	6	1.3	244	2	T26265	hypothetical prote	825	6	1.3	261	2	T11540	cytochrome-c oxida
753	6	1.3	245	1	KYBOA	chymotrypsin (EC 3	826	6	1.3	261	2	T11306	cytochrome-c oxida
754	6	1.3	245	2	T31187	transmembrane prote	827	6	1.3	262	2	C84776	hypothetical prote
755	6	1.3	245	2	C97006	transmembrane prote	828	6	1.3	262	2	T11883	cytochrome-c oxida
756	6	1.3	245	2	H97438	hypothetical prote	829	6	1.3	262	2	F87498	exodeoxyribonuclea
757	6	1.3	246	2	S38344	CDEr-binding prote	830	6	1.3	263	2	H84918	glutathione S-tran
758	6	1.3	246	2	S47805	hypothetical 27.4k	831	6	1.3	263	2	T18880	hypothetical prote
759	6	1.3	246	2	D91186	probable outer mem	832	6	1.3	263	2	AD2114	hypothetical prote

833	6	1.3	263	2	T02293	hypothetical prote	906	6	1.3	281	2	A72561	hypothetical prote
834	6	1.3	263	2	H86147	TiN6.7 protein - A	907	6	1.3	281	2	T20894	hypothetical prote
835	6	1.3	264	2	D81253	probable peptidase A	908	6	1.3	281	2	T32765	hypothetical prote
836	6	1.3	264	2	S78700	probable export pr	909	6	1.3	282	2	T14225	NADH2 dehydrogenas
837	6	1.3	264	2	AF0753	flagellar biosynth	910	6	1.3	282	2	AE1559	glycine betaine AB
838	6	1.3	264	2	D71197	hypothetical prote	911	6	1.3	282	2	AG1201	glycine betaine AB
839	6	1.3	264	2	D85097	hypothetical prote	912	6	1.3	282	2	D72771	probable bacterioc
840	6	1.3	264	2	TJ00328	cellulase (EC 3.2.	913	6	1.3	282	2	T19594	hypothetical prote
841	6	1.3	264	2	T34320	hypothetical prote	914	6	1.3	282	2	T03906	cytochrome B561 ho
842	6	1.3	265	2	AB2000	phosphatidate cyti	915	6	1.3	283	2	A69207	dhvdrodipicolinat
843	6	1.3	265	2	B81229	phosphatidate cyti	916	6	1.3	283	2	D81199	conserved hypotnet
844	6	1.3	265	2	C82358	glutamate racemase	917	6	1.3	284	2	D81834	conserved hypotnet
845	6	1.3	265	2	A41327	dechloromethane de	918	6	1.3	284	2	F89004	protein T24A6.13 l
846	6	1.3	265	2	A75158	hypothetical prote	919	6	1.3	284	2	C86477	protein F1504.21 l
847	6	1.3	265	2	T46089	proline-rich prote	920	6	1.3	284	2	E97054	sporulation proteol
848	6	1.3	266	1	ELR11	pancreatic elastas	921	6	1.3	285	2	B87426	rhodanese family p
849	6	1.3	266	1	ELVUF	pancreatic elastas	922	6	1.3	285	2	G02240	conserved hypotnet
850	6	1.3	267	1	TVHUF5	fibroblast growth	923	6	1.3	285	2	E69905	transcription regu
851	6	1.3	267	2	AB7404	ribosomal protein	924	6	1.3	285	2	D95276	probable ABC trans
852	6	1.3	267	2	JN0831	GlnR protein - Str	925	6	1.3	286	2	H89003	protein T24A6.5 l
853	6	1.3	267	4	A56615	probable pancreati	926	6	1.3	286	2	AG2000	hypothetical prote
854	6	1.3	268	2	D86270	F21F23.14 protein	927	6	1.3	286	2	AD3263	hypothetical prote
855	6	1.3	268	2	AB1457	high-affinity zinc	928	6	1.3	287	2	AF0679	ribonuclease bn (E
856	6	1.3	268	2	AD1094	high-affinity zinc	929	6	1.3	287	2	T51706	starvation sensing
857	6	1.3	268	2	C90566	hypothetical prote	930	6	1.3	287	2	E70925	formyltetrahydrofo
858	6	1.3	269	2	AB6889	zinc ABC transport	931	6	1.3	287	2	C98102	probable ribosomal
859	6	1.3	269	2	E87707	ABC transporter, p	932	6	1.3	287	2	AD3150	hexulose-6-phospha
860	6	1.3	269	2	AF0134	flagellar biosynth	933	6	1.3	288	2	AB4058	transcription regu
861	6	1.3	270	2	F97433	putrescine transpo	934	6	1.3	289	2	S17725	formyltetrahydrofo
862	6	1.3	270	2	AG2651	ABC transporter, m	935	6	1.3	289	2	AB3608	H+-transporting tw
863	6	1.3	270	2	T66118	hypothetical prote	936	6	1.3	289	2	F70845	polymaine transpor
864	6	1.3	270	2	E84578	probable senescenc	937	6	1.3	289	2	AH2602	hypothetical prote
865	6	1.3	270	2	F84293	hypothetical prote	938	6	1.3	289	2	H97384	conserved hypotnet
866	6	1.3	270	2	A75064	hypothetical prote	939	6	1.3	290	2	A36134	hypothetical prote
867	6	1.3	270	2	C88483	protein C05D11.13	940	6	1.3	290	2	F70458	Repa protein - Esc
868	6	1.3	271	2	D83059	phosphatidylserine	941	6	1.3	291	2	AG1115	hypothetical prote
869	6	1.3	271	2	D85035	hypothetical prote	942	6	1.3	291	2	T04494	transcription regu
870	6	1.3	272	2	AC2431	permease protein o	943	6	1.3	291	2	B81130	hypothetical prote
871	6	1.3	272	2	JC7357	3IK proliferation	944	6	1.3	291	2	H81836	conserved hypotnet
872	6	1.3	273	1	B64623	lipopolysaccharide	945	6	1.3	292	2	T40637	probable integral
873	6	1.3	273	2	S39200	formimidopyrimidin	946	6	1.3	292	2	G97111	40s ribosomal prot
874	6	1.3	273	2	C71369	probable sugar ABC	947	6	1.3	292	2	D91229	uncharacterized st
875	6	1.3	273	2	AH3303	putrescine transpo	948	6	1.3	292	2	C60706	probable aldolase
876	6	1.3	273	2	S00953	gag protein - frui	949	6	1.3	292	2	S40825	hypothetical 32k p
877	6	1.3	273	2	D70428	hypothetical prote	950	6	1.3	293	2	AB4110	sugar ABC transpor
878	6	1.3	274	2	G95345	conserved hypotnet	951	6	1.3	293	2	AH2306	conserved hypotnet
879	6	1.3	274	2	F90351	hypothetical prote	952	6	1.3	293	2	AH0017	heat-shock chapero
880	6	1.3	275	2	T22414	hypothetical prote	953	6	1.3	293	2	T04450	hypothetical prote
881	6	1.3	276	2	C83009	probable ATP-blind	954	6	1.3	293	2	AB5817	hypothetical prote
882	6	1.3	276	2	T15455	hypothetical prote	955	6	1.3	293	2	E85718	hypothetical prote
883	6	1.3	276	2	E86756	prophage p12 prote	956	6	1.3	294	2	H97655	probable tail comp
884	6	1.3	277	1	QREEST	sulfate/chiosulfat	957	6	1.3	294	2	AE2879	purU-like protein
885	6	1.3	277	2	B85885	sulfate transport	958	6	1.3	294	2	AB3558	formyltetrahydrofo
886	6	1.3	277	2	G91040	sulfate transport	959	6	1.3	294	2	T34199	formyltetrahydrofo
887	6	1.3	277	2	AH0811	hypothetical prote	960	6	1.3	294	2	T41953	G protein-coupled
888	6	1.3	277	2	G71456	hypothetical prote	961	6	1.3	294	2	AD1255	protein secretion
889	6	1.3	277	2	SI6563	noIG protein - Rhi	962	6	1.3	294	2	AI1617	protein secretion
890	6	1.3	277	2	AF2404	hypothetical prote	963	6	1.3	295	2	B83587	probable polynucle
891	6	1.3	278	2	AS6421	casein kinase II (	964	6	1.3	295	2	JE0174	frizled protein-2
892	6	1.3	278	2	T21718	hypothetical prote	965	6	1.3	295	2	T43463	hypothetical prote
893	6	1.3	278	2	E83453	hypothetical prote	966	6	1.3	296	2	T06572	convicillin precurs
894	6	1.3	278	2	T46458	hypothetical prote	967	6	1.3	296	2	AF1170	probable sugar ABC
895	6	1.3	278	2	D71256	conserved hypotnet	968	6	1.3	296	2	AH1527	probable sugar ABC
896	6	1.3	278	2	F75109	hypothetical prote	969	6	1.3	296	2	D98345	SN-glycerol-3-phos
897	6	1.3	279	2	JC2106	tolB protein - Agr	970	6	1.3	296	2	AC2937	hypothetical prote
898	6	1.3	279	2	S52639	hypothetical prote	971	6	1.3	296	2	T30575	hypothetical prote
899	6	1.3	280	2	C70696	probable transpor	972	6	1.3	296	2	F81155	conserved hypotnet
900	6	1.3	280	2	F96768	hypothetical prote	973	6	1.3	296	2	H82977	4-hydroxybenzoate-
901	6	1.3	281	2	D45313	putrescine transpo	974	6	1.3	297	2	EB6701	hypothetical prote
902	6	1.3	281	2	A90746	permease of putres	975	6	1.3	298	2	E91229	probable dehydroge
903	6	1.3	281	2	E85596	putrescine transpo	976	6	1.3	298	2	D86076	probable dehydroge
904	6	1.3	281	2	AI0162	putrescine transpo	977	6	1.3	298	2	AI0947	probable oxidoredu
905	6	1.3	281	2	AE0606	putrescine transpo	978	6	1.3	298	2	S40826	hypothetical 31.2K

```

979 6 1.3 298 2 E48213 syntaxin 4 - rat
980 6 1.3 298 2 T22264 hypothetical prote
981 6 1.3 298 2 E83101 hypothetical prote
982 6 1.3 298 2 E83783 Mn catalase BH1069
983 6 1.3 298 2 E95286 hypothetical prote
984 6 1.3 299 2 S59144 NADH2 dehydrogenas
985 6 1.3 299 2 AC3093 transcription regu
986 6 1.3 300 2 C69857 formyltetrahydrofo
987 6 1.3 300 2 T46751 inner membrane pro
988 6 1.3 300 2 T49656 hypothetical prote
989 6 1.3 300 2 F97623 hypothetical prote
990 6 1.3 300 2 AF2846 permease [imported
991 6 1.3 301 1 A41894 fructokinase (EC 2
992 6 1.3 301 2 H70505 hypothetical prote
993 6 1.3 301 2 A12483 hypothetical prote
994 6 1.3 301 2 B84254 hypothetical prote
995 6 1.3 301 2 T37031 hypothetical prote
996 6 1.3 302 2 C64103 glycine-tRNA ligas
997 6 1.3 302 2 A45511 chitinase (EC 3.2.
998 6 1.3 302 2 D82112 tyra protein VC214
999 6 1.3 302 2 T50784 hypothetical prote
1000 6 1.3 303 1 SYECGA glycine-tRNA ligas

```

## ALIGNMENTS

## RESULT 1

```

S58396
presentin 1, splice form 467 - human
N:Alternate names: Alzheimer's disease protein 3; protein S182
C:Species: Homo sapiens (man)
C>Date: 29-Jan-1998 #sequence_revision 13-Feb-1998 #text_change 29-Sep-1999
C:Accession: S58396; S71401; S71402
R:Sherrington, R.; Rogeev, E.I.; Liang, Y.; Rogeeva, E.A.; Levesque, G.; Ikeda, M.; Chi,
ero, I.; Pinessi, L.; Nee, L.; Chumakov, I.; Pollen, D.; Brookes, A.; Sansseau, P.; Polin
E.; Rommens, J.M.; St George-Hyslop, P.H.
Nature 375, 754-760, 1995
A:title: Cloning of a gene bearing missense mutations in early-onset familial Alzheimer'
A:Reference number: 158095; MUID:95319502; PMID:7596406
A:Accession: S58396
A:Molecule type: mRNA
A:Residues: 1-467 <SHE>
A:Cross-references: EMBL:142110; NID:9904118; PIDN:AAB46416.1; PID:9904119
R:Vidal, R.; Ghiso, J.; Wisniewski, T.; Frangione, B.
FEBS Lett. 393, 19-23, 1996
A:title: Alzheimer's presentin 1 gene expression in platelets and megakaryocytes. Ident
A:Reference number: S71401; MUID:96397521; PMID:8804415
A:Accession: S71401
A:Molecule type: mRNA
A:Status: not compared with conceptual translation
A:Residues: 24-32;254-256, 290-292;316-317,376-379 <VIW>
A:Experimental source: Dami megakaryotic cell line (ATCC CRL-9792) and platelets
C:Genetics:
A:Gene: GDB:PSEN1; AD3; PAD; S182; PS1
A:Cross-references: GDB:135682; OMIM:104311
A:Map position: 14q24.3-14q24.3
C:Superfamily: presentin
C:Keywords: alternative splicing; Alzheimer's disease; glycoprotein; transmembrane prote

```

June 29

```

F:82-100/Domain: transmembrane #status predicted <TM1>
F:133-154/Domain: transmembrane #status predicted <TM2>
F:164-185/Domain: transmembrane #status predicted <TM3>
F:195-213/Domain: transmembrane #status predicted <TM4>
F:221-238/Domain: transmembrane #status predicted <TM5>
F:244-264/Domain: transmembrane #status predicted <TM6>
F:281-301/Domain: transmembrane #status predicted <TM7>
F:408-428/Domain: transmembrane #status predicted <TM8>
F:433-453/Domain: transmembrane #status predicted <TM9>
F:279,405/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

Query Match 78.4%; Score 366; DB 2; Length 467;

Best Local Similarity 99.8%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;

Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MTELPAPLSYFQNAQSEDNHLSNTVRSQNDNREQEHNDRLSGHPEPLSNGRPGNSR 60
|||||
Db 1 MTELPAPLSYFQNAQSEDNHLSNTVRSQNDNREQEHNDRLSGHPEPLSNGRPGNSR 60
QY 61 QVVEQDEDEDELTLYGAKHVMLEVPVTLCAVNVVATKSVSEYTRKDGOLYTPFTE 120
|||||
Db 61 QVVEQDEDEDELTLYGAKHVMLEVPVTLCAVNVVATKSVSEYTRKDGOLYTPFTE 120
QY 121 DTEVVGORALHSTILNAIMISVIVVMTLLVLYKYKCYVIAHMLTSSLLFFPSFI 180
|||||
Db 121 DTEVVGORALHSTILNAIMISVIVVMTLLVLYKYKCYVIAHMLTSSLLFFPSFI 180
QY 181 YLGEVFETVYVAVDYITVALLINMGVGMISIHMKPRLQQAYLIMISALMALVIRY 240
|||||
Db 181 YLGEVFETVYVAVDYITVALLINMGVGMISIHMKPRLQQAYLIMISALMALVIRY 240
QY 241 LPEWTAMLIAVISVVDVAVLCPKPLRMIVETAQRNETLFPALYSSTMVLYVMAE 300
|||||
Db 241 LPEWTAMLIAVISVVDVAVLCPKPLRMIVETAQRNETLFPALYSSTMVLYVMAE 300
QY 301 GDPEAQRVSKNSKYNAESTERESQDTVAENDGCFSEMEADQSDHLSGRSTPESRAA 360
|||||
Db 301 GDPEAQRVSKNSKYNAESTERESQDTVAENDGCFSEMEADQSDHLSGRSTPESRAA 360
QY 361 VOELSSILAGEDPEERGVRLGLGDFIFYSVLVGKASATASGDMNTTACFVAIILGLCL 420
|||||
Db 361 VOELSSILAGEDPEERGVRLGLGDFIFYSVLVGKASATASGDMNTTACFVAIILGLCL 420
QY 421 TLLILAIFKKALPALPISITFGVLFVFRADYLVQPFMDQAFHQFYI 467
|||||
Db 421 TLLILAIFKKALPALPISITFGVLFVFRADYLVQPFMDQAFHQFYI 467

```

## RESULT 2

```

S63683
presentin 1-463 - human
C:Species: Homo sapiens (man)
C>Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 29-Sep-1999
C:Accession: S63683
R:Sahara, N.; Yahagi, Y.; Takagi, H.; Kondo, T.; Okochi, M.; Usami, M.; Shirasawa, T.
FEBS Lett. 381, 7-11, 1996
A:title: Identification and characterization of presentin 1-463, I-463 and I-374.
A:Reference number: S63683; MUID:96193901; PMID:8641442
A:Accession: S63683
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-463 <SAH>
A:Cross-references: EMBL:U40379; NID:91244637; PIDN:AAB05894.1; PID:91244638
C:Superfamily: presentin

```

Query Match 72.2%; Score 337; DB 2; Length 463;

Best Local Similarity 99.8%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;

Matches 437; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 30 NDRERQEHNDRLSLGHPEPLSNGRPGNSQVVEQDEDEDELTLYGAKHVMLEVPV 89
|||||
Db 26 NDRERQEHNDRLSLGHPEPLSNGRPGNSQVVEQDEDEDELTLYGAKHVMLEVPV 85
QY 90 TLLCMVVVATIKSVSFTRKDGOLYTPFTEDETVGQALHSILNAIMISVIVVMTIL 149
|||||
Db 86 TLLCMVVVATIKSVSFTRKDGOLYTPFTEDETVGQALHSILNAIMISVIVVMTIL 145
QY 150 LVVLVYRCVKVIAHMLTSSLLFFPSFIYGEVFKTYNAVVDYITVALLIWNGVVG 209
|||||
Db 146 LVVLVYRCVKVIAHMLTSSLLFFPSFIYGEVFKTYNAVVDYITVALLIWNGVVG 205
QY 210 MISIHMKPRLQQAYLIMISALMALVFIKYLEPWTAMLIAVISVVDVAVLCPKPLR 269
|||||
Db 206 MISIHMKPRLQQAYLIMISALMALVFIKYLEPWTAMLIAVISVVDVAVLCPKPLR 265
QY 270 MLVETAQRNETLFPALYSSTMVLYVMAEGDPEAQRVSKNSKYNAESTERESQDTVA 329
|||||

```

Db 266 MLVETAOERNETLFPALITYSTWVWLVNMAEGDPEAQRVSKSKYNAESTRESQDTVA 325  
Qy 330 ENDGCGFSEEREAQRDHLGPHRSPESRAVQELSSSLAGEDEEGVNLGLCDFEY 389  
Db 326 ENDGCGFSEEREAQRDHLGPHRSPESRAVQELSSSLAGEDEEGVNLGLCDFEY 385  
Qy 390 SVLVGKASATASCDMNTTIFACFVAILIGLCTLLLLAFKKALPALPISIFGLVEFYAT 449  
Db 386 SVLVGKASATASCDMNTTIFACFVAILIGLCTLLLLAFKKALPALPISIFGLVEFYAT 445  
Qy 450 DYLVOPEMDQLAFHQFYI 467  
Db 446 DYLVOPEMDQLAFHQFYI 463  
RESULT 3  
presenilin 1, splice form 374 - human  
N:Alternate names: Alzheimer's disease protein 3  
C:Species: Homo sapiens (man)  
C>Date: 20-Jul-1996 #sequence\_revision 13-Mar-1997 #text\_change 29-Sep-1999  
C:Accession: S63684  
R:Shahar, N.; Yahagi, Y.; Takagi, H.; Kondo, T.; Okochi, M.; Usami, M.; Shirasawa, T.; M  
FEBS Lett. 381, 7-11, 1996  
A:Title: Identification and characterization of presenilin I-467, I-463 and I-374.  
A:Reference number: S63683; MUID:96193901; PMID:8641442  
A:Accession: S63684  
A:Status: preliminary; nucleic acid sequence not shown  
A:Residues: 1-374 <SAH>  
A:Molecule type: mRNA  
A:Cross-references: EMBL:U40380; NID:g1244639; PID:AA05895.1; PID:g1244640  
C:Genetics:  
A:Gene: GDB:PSEN1; AD3; FAD; S182; PS1  
A:Cross-references: GDB:135682; OMIM:104311  
A:Map position: 14q24.3-14q24.3  
C:Superfamily: presenilin  
Query Match 40.3%; Score 188; DB 2; Length 374;  
Best Local Similarity 99.7%; Pred. No. 4e-175;  
Matches 288; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 30 NDNREQEHNDRRSLGHEPLSNRPGNSROVVEQDEDEDELTKYGAKEHMLFVYV 89  
Db 26 NDNREQEHNDRRSLGHEPLSNRPGNSROVVEQDEDEDELTKYGAKEHMLFVYV 85  
Qy 90 TLCWVNVVATIKTSVSFYTRKDGQLITPTEDTETVGGRAHSLINAAIMSVIVMTIL 149  
Db 86 TLCWVNVVATIKTSVSFYTRKDGQLITPTEDTETVGGRAHSLINAAIMSVIVMTIL 145  
Qy 150 LVVLKYKCYVTHAMLLTSSLLFFPSFTYLGVEFTYNAVYITVALLIMLVGVG 209  
Db 146 LVVLKYKCYVTHAMLLTSSLLFFPSFTYLGVEFTYNAVYITVALLIMLVGVG 205  
Qy 210 MISIMKGPLRLOQAYLIMISALMALVFIKYLEPTAMLLAVISVYDLVAVLCRGLR 269  
Db 206 MISIMKGPLRLOQAYLIMISALMALVFIKYLEPTAMLLAVISVYDLVAVLCRGLR 265  
Qy 270 MLVETAOERNETLFPALITYSTWVWLVNMAEGDPEAQRVSKSKYNAE 318  
Db 266 MLVETAOERNETLFPALITYSTWVWLVNMAEGDPEAQRVSKSKYNAE 314  
RESULT 4  
JC5081  
presenilin 1 protein isoform 463 - lesser mouse lemur  
C:Species: Microcebus murinus (lesser mouse lemur)  
C>Date: 31-Jan-1997 #sequence\_revision 31-Jan-1997 #text\_change 13-Sep-1998  
C:Accession: JC5081  
R:Calenda, A.; Mestire-Frances, N.; Czech, C.; Pradler, L.; Pelter, A.; Bons, N.; Bellis,  
Biochem. Biophys. Res. Commun. 228, 430-439, 1996  
A:Title: Molecular cloning, sequencing, and brain expression of the presenilin 1 gene in  
A:Reference number: JC5080; MUID:97079199; PMID:8920931  
A:Contents: brain

A:Accession: JC5081  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-463 <CAL>  
A:Cross-references: EMBL:Z71333  
C:Comment: This protein is an intermembrane protein with seven transmembrane domains.  
C:Genetics:  
A:Gene: ps1  
A:Map position: 14  
C:Superfamily: presenilin  
C:Keywords: transmembrane protein  
F:78-96/Domain: transmembrane #status predicted <TM1>  
F:129-150/Domain: transmembrane #status predicted <TM2>  
F:160-181/Domain: transmembrane #status predicted <TM3>  
F:191-209/Domain: transmembrane #status predicted <TM4>  
F:217-234/Domain: transmembrane #status predicted <TM5>  
F:240-257/Domain: transmembrane #status predicted <TM6>  
F:404-424/Domain: transmembrane #status predicted <TM7>  
Query Match 22.9%; Score 107; DB 2; Length 463;  
Best Local Similarity 100.0%; Pred. No. 5.3e-96;  
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 206 GVVGMISIMKGPLRLOQAYLIMISALMALVFIKYLEPTAMLLAVISVYDLVAVLCR 265  
Db 202 GVVGMISIMKGPLRLOQAYLIMISALMALVFIKYLEPTAMLLAVISVYDLVAVLCR 261  
Qy 266 GPLRLVETAOERNETLFPALITYSTWVWLVNMAEGDPEAQRVSKN 312  
Db 262 GPLRLVETAOERNETLFPALITYSTWVWLVNMAEGDPEAQRVSKN 308  
RESULT 5  
JC5080  
presenilin 1 protein isoform 467 - lesser mouse lemur  
C:Species: Microcebus murinus (lesser mouse lemur)  
C>Date: 31-Jan-1997 #sequence\_revision 31-Jan-1997 #text\_change 20-Jun-2000  
C:Accession: JC5080  
R:Calenda, A.; Mestire-Frances, N.; Czech, C.; Pradler, L.; Pelter, A.; Bons, N.; Bell  
Biochem. Biophys. Res. Commun. 228, 430-439, 1996  
A:Title: Molecular cloning, sequencing, and brain expression of the presenilin 1 gene  
A:Reference number: JC5080; MUID:97079199; PMID:8920931  
A:Accession: JC5080  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-467 <CAL>  
A:Cross-references: EMBL:Z71333; NID:g1707591; PID:CAA95930.1; PID:g1707592  
A:Experimental source: brain  
C:Comment: This protein is an intergal membrane protein with seven transmembrane dom  
C:Genetics:  
A:Gene: ps1  
A:Map position: 14  
C:Superfamily: presenilin  
C:Keywords: transmembrane protein  
F:82-100/Domain: transmembrane #status predicted <TM1>  
F:133-154/Domain: transmembrane #status predicted <TM2>  
F:164-185/Domain: transmembrane #status predicted <TM3>  
F:195-213/Domain: transmembrane #status predicted <TM4>  
F:221-238/Domain: transmembrane #status predicted <TM5>  
F:244-261/Domain: transmembrane #status predicted <TM6>  
F:408-428/Domain: transmembrane #status predicted <TM7>  
Query Match 22.9%; Score 107; DB 2; Length 467;  
Best Local Similarity 100.0%; Pred. No. 5.4e-96;  
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 206 GVVGMISIMKGPLRLOQAYLIMISALMALVFIKYLEPTAMLLAVISVYDLVAVLCR 265  
Db 206 GVVGMISIMKGPLRLOQAYLIMISALMALVFIKYLEPTAMLLAVISVYDLVAVLCR 265  
Qy 266 GPLRLVETAOERNETLFPALITYSTWVWLVNMAEGDPEAQRVSKN 312  
Db 266 GPLRLVETAOERNETLFPALITYSTWVWLVNMAEGDPEAQRVSKN 312

## RESULT 6

I78388

S182 protein - mouse

C:Species: Mus musculus (house mouse)

C&gt;Date: 27-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 29-Sep-1999

C:Accession: I78388

A:Title: Identification and expression analysis of a potential familial Alzheimer's d

A:Reference number: I39174; MUID:96109229; PMID:8618867

A:Accession: I39174

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-442 &lt;RES&gt;

A:Cross-references: EMBL:U34349; NID:91079575; PIDN:AAC50290.1; PID:91079576

C:Superfamily: presentin

## Query Match

20.8%; Score 97; DB 2; Length 467;

Best Local Similarity 100.0%; Pred. No. 3.1e-86;

Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 213 IHMKGPLRLOQAVLIMISALMALVFIRYLPMTMLLAVISYDYLAVLCPRKPLRLMLV 272

DB 213 IHMKGPLRLOQAVLIMISALMALVFIRYLPMTMLLAVISYDYLAVLCPRKPLRLMLV 272

Query 273 ETAOERNETLFPALISYSTMVWLVNMAEGDPEAQRV 309

DB 273 ETAOERNETLFPALISYSTMVWLVNMAEGDPEAQRV 309

## RESULT 7

JCS5390

presentin-alpha - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C&gt;Date: 04-Jun-1997 #sequence\_revision 18-Jul-1997 #text\_change 20-Jun-2000

C:Accession: JCS5390

A:Title: Cloning of Xenopus presentin-alpha and -beta cDNAs and their differential exp

A:Reference number: JCS5390; MUID:97223465; PMID:9070286

A:Accession: JCS5390

A&gt;Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-433 &lt;RSU&gt;

A:Cross-references: DDBJ:D84427; NID:91944353; PIDN:BA19570.1; PID:91944354

A:Experimental source: brain

C:Comment: This protein plays a role in negative regulation of apoptotic cascades during

C:Superfamily: presentin

F:48-66/Domain: transmembrane #status predicted &lt;TM1&gt;

F:99-119/Domain: transmembrane #status predicted &lt;TM2&gt;

F:130-149/Domain: transmembrane #status predicted &lt;TM3&gt;

F:161-178/Domain: transmembrane #status predicted &lt;TM4&gt;

F:187-203/Domain: transmembrane #status predicted &lt;TM5&gt;

F:210-227/Domain: transmembrane #status predicted &lt;TM6&gt;

F:374-394/Domain: transmembrane #status predicted &lt;TM7&gt;

Query Match 12.6%; Score 59; DB 2; Length 433;

Best Local Similarity 100.0%; Pred. No. 3.4e-49;

Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 409 ACPVALIGLCTLLLAIFKKAIPALPISITRGVYFATDYIVQFMQALFHQYI 467

DB 375 ACPVALIGLCTLLLAIFKKAIPALPISITRGVYFATDYIVQFMQALFHQYI 433

## RESULT 8

I39174

seven trans-membrane domain protein AD3LP/AD5 - human

C:Species: Homo sapiens (man)

C&gt;Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 29-Sep-1999

C:Accession: I39174

A:Title: Identification and expression analysis of a potential familial Alzheimer's d

A:Reference number: I39174; MUID:96109229; PMID:8618867

A:Accession: I39174

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-442 &lt;RES&gt;

A:Cross-references: EMBL:U34349; NID:91079575; PIDN:AAC50290.1; PID:91079576

C:Superfamily: presentin

## Query Match

6.0%; Score 28; DB 2; Length 442;

Best Local Similarity 100.0%; Pred. No. 6.2e-19;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 253 ISYVDLVAVLCPRKGPLRLMLVETAOERNE 280

DB 259 ISYVDLVAVLCPRKGPLRLMLVETAOERNE 286

## RESULT 9

A56993

presentin 2 - human

N:Alternate names: Alzheimer's disease protein 4

C:Species: Homo sapiens (man)

C&gt;Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 17-Nov-2000

C:Accession: A56993; I58098

A:Title: Candidate gene for the chromosome 1 familial Alzheimer's disease locus.

A:Reference number: A56993; MUID:95365816; PMID:7638652

A:Accession: A56993

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-448 &lt;RES&gt;

A:Cross-references: GB:L43964; NID:951202; PIDN:AAB59557.1; PID:951203

A:Experimental source: R:Rogaev, E.I.; Sherrington, R.; Rogaeva, E.A.; Levesque, G.; Ikeda, M.; Liang, Y.; C

; Cohen, D.; Tanfild, L.; Fraser, P.E.; Rommens, J.M.; St George-Hyslop, P.H.

Nature 376, 775-778, 1995

A:Title: Familial Alzheimer's disease in kindreds with missense mutations in a gene o

A:Reference number: I58098; MUID:95379971; PMID:7651536

A:Accession: I58098

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-122 'P', 124-448 &lt;RE2&gt;

A:Cross-references: GB:L44577; NID:9950347; PIDN:AAC42012.1; PID:9950348

A:Gene: GDB:PSEN2; AD4; STM2; PS2; ES-1

A:Cross-references: GDB:633044; OMIM:600759

A:Map position: 1q31-1q42

C:Superfamily: presentin

## Query Match

6.0%; Score 28; DB 2; Length 448;

Best Local Similarity 100.0%; Pred. No. 6.3e-19;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 253 ISYVDLVAVLCPRKGPLRLMLVETAOERNE 280

DB 259 ISYVDLVAVLCPRKGPLRLMLVETAOERNE 286

## RESULT 10

JCS391

presentin-beta - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C&gt;Date: 04-Jun-1997 #sequence\_revision 18-Jul-1997 #text\_change 20-Jun-2000

C:Accession: JCS391

A:Title: Cloning of Xenopus presentin-alpha and -beta cDNAs and their differential e

A:Reference number: JCS390; MUID:97223465; PMID:9070286

A:Accession: J05391  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1449 <TSU>  
 A:Cross-references: DDBJ:D84428; NID:g1944355; PIDN:BAI9571.1; PID:g1944356  
 A:Experimental source: brain  
 C:Comment: This protein plays a role in negative regulation of apoptotic cascades during  
 C:Superfamily: presenilin  
 F:91-109/Domain: transmembrane #status predicted <TM1>  
 F:142-162/Domain: transmembrane #status predicted <TM2>  
 F:173-192/Domain: transmembrane #status predicted <TM3>  
 F:204-221/Domain: transmembrane #status predicted <TM4>  
 F:230-246/Domain: transmembrane #status predicted <TM5>  
 F:252-279/Domain: transmembrane #status predicted <TM6>  
 F:390-410/Domain: transmembrane #status predicted <TM7>

Query Match 5.8%; Score 27; DB 2; Length 449;  
 Best Local Similarity 100.0%; Pred. No. 6e-18;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 81 HVIMLFVPTLCMVVVATKSVSYFT 107  
 Db 90 HVIMLFVPTLCMVVVATKSVSYFT 116

RESULT 11  
 S60253  
 sel-12 protein - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 10-Apr-1996 #sequence\_revision 19-Apr-1996 #text\_change 13-Sep-1998  
 C:Accession: S60253  
 R:Levitan, D.; Greenwald, I.  
 Nature 377, 351-354, 1995  
 A:Title: Facilitation of lin-12-mediated signalling by sel-12, a Caenorhabditis elegans  
 A:Reference number: S60253; MID:96032531; PMID:7566091  
 A:Accession: S60253  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1441 <LEV>  
 A:Cross-references: EMBL:U35560  
 C:Superfamily: Presenilin

Query Match 3.9%; Score 18; DB 2; Length 461;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-09;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 375 EERGVKLGDFIFYSVL 392  
 Db 354 EERGVKLGDFIFYSVL 371

RESULT 12  
 E89453  
 Protein F35H12.3 [imported] - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Nov-2001  
 C:Accession: E89453  
 R:anonymous, The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
 A:Reference number: A75000; MID:99069613; PMID:9851916  
 A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C\_ele  
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A:Accession: E89453  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-836 <STO>  
 A:Cross-references: GB:chr.X; PIDN:AAA8176.1; PID:q1109863; GSPDB:GN00028; CESP:F35H12.  
 A:Note: similar to M. musculus Sl82 protein (GB:I42177)  
 C:Genetics:  
 A:Gene: F35H12.3  
 A:Map position: X  
 A:Map position: X

Query Match 3.9%; Score 18; DB 2; Length 836;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-09;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 375 EERGVKLGDFIFYSVL 392  
 Db 354 EERGVKLGDFIFYSVL 371

RESULT 13  
 A84702  
 Probable presenilin [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
 C:Accession: A84702  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon,  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MID:20083487; PMID:10617197  
 A:Accession: A84702  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-397 <STO>  
 A:Cross-references: GB:AE002093; NID:g4567215; PIDN:AAD23630.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g29900  
 A:Map position: 2  
 C:Superfamily: presenilin

Query Match 3.2%; Score 15; DB 2; Length 397;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 380 KLGIDGDFIFYSVLG 394  
 Db 313 KLGIDGDFIFYSVLG 327

RESULT 14  
 T00724  
 presenilin homolog F22013.19 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 22-Oct-1999  
 C:Accession: T00724  
 R:Shim, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.;  
 eologists, A.; Ecker, J.R.  
 submitted to the EMBL Data Library, April 1998  
 A:Description: Genomic sequence for Arabidopsis thaliana BAC F22013.  
 A:Reference number: Z14200  
 A:Accession: T00724  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-453 <SHI>  
 A:Cross-references: EMBL:AC003981; NID:g3063438; PID:g3063457; GSPDB:GN00059; ATSP:F2  
 A:Experimental source: cultivar Columbia  
 C:Genetics:  
 A:Gene: ATSP:F22013.19  
 A:Map position: 1  
 A:Introns: 108/1  
 C:Superfamily: presenilin

Query Match 3.2%; Score 15; DB 2; Length 453;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 380 KLGIDGDFIFYSVLG 394  
 Db 369 KLGIDGDFIFYSVLG 383

RESULT 15

G90492  
hypothetical protein SSO3092 [imported] - Sulfolobus solfataricus  
C:Species: Sulfolobus solfataricus  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
C:Accession: G90492  
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan-  
'Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.  
arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.  
A:Description: Sulfolobus solfataricus complete genome.  
A:Reference number: A99139  
A:Accession: G90492  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-334 <KUR>  
A:Cross-references: GB:AE006641; NID:q13816512; PIDN:AAK43198.1; GSPDB:GN00155  
C:Genetics: SSO3092

Query Match 1.9%; Score 9; DB 2; Length 334;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 172 LLLFFSFI 180  
|||||||  
DB 198 LLLFFSFI 206

RESULT 16  
H65055  
hypothetical protein b2748 - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
C:Accession: H65055  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co-  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: H65055  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-103 <BLAT>  
A:Cross-references: GB:AE000358; GB:U00096; NID:q2367156; PIDN:AACT5790.1; PID:91789105;  
A:Experimental source: strain K-12, substrain MG1655  
C:Superfamily: hypothetical protein HI0673

Query Match 1.7%; Score 8; DB 2; Length 103;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 420 LTTLLAI 427  
|||||||  
DB 4 LTTLLAI 11

RESULT 17  
B91079  
hypothetical protein ECS3602 [imported] - Escherichia coli (strain O157:H7, substrain R1)  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
C:Accession: B91079  
R:Havashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno-  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: B91079  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-103 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA037025.1; PID:q13363073; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain R1MD 0509952

C:Genetics:  
A:Gene: ECS3602  
C:Superfamily: hypothetical protein HI0673

Query Match 1.7%; Score 8; DB 2; Length 103;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 420 LTTLLAI 427  
|||||||  
DB 4 LTTLLAI 11

RESULT 18  
C85924  
hypothetical protein Z4056 [imported] - Escherichia coli (strain O157:H7, substrain E)  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: C85924  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May-  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apoda-  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: C85924  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-103 <STO>  
A:Cross-references: GB:AE005174; NID:q12517202; PIDN:AACT57855.1; GSPDB:GN00145; UMGCP:  
A:Experimental source: strain O157:H7, substrain ED1933  
C:Genetics:  
A:Gene: Z4056  
C:Superfamily: hypothetical protein HI0673

Query Match 1.7%; Score 8; DB 2; Length 103;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 420 LTTLLAI 427  
|||||||  
DB 4 LTTLLAI 11

RESULT 19  
A49304  
ventral prostate major secretory protein, 20K - rat  
M:Alternate names: 22K glycoprotein, prostatic; cystatin-related protein 1  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 06-Oct-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Apr-2001  
C:Accession: A49304; S29944; A32819; J06615; A34552; A49200; S38813  
R:Ho, K.C.; Marschke, K.B.; Tan, J.; Power, S.G.; Wilson, E.M.; French, F.S.  
J. Biol. Chem. 268, 27226-27235, 1993  
A:Title: A complex response element in intron 1 of the androgen-regulated 20-kDa prot-  
A:Reference number: A49304; MUID:94086539; PMID:8262963  
A:Accession: A49304  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-176 <HO1>  
A:Cross-references: GB:U12454; NID:q499666; PIDN:AAA40732.1; PID:q499667  
A:Note: sequence extracted from NCBI backbone (NCBIN:140973, NCBI:140974)  
R:Peeters, B.  
submitted to the EMBL Data Library, July 1992  
A:Reference number: S29944  
A:Accession: S29944  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-176 <PEE>  
A:Cross-references: EMBL:Z13993; NID:q56965; PIDN:CA078384.1; PID:q56966  
R:Ho, K.C.; Snook, R.; Quarumby, V.; Viskochil, D.H.; Rennie, P.S.; Wilson, E.M.; Fren-  
Biochemistry 28, 6367-6373, 1989  
A:Title: Primary structure and androgen regulation of a 20-kilodalton protein specific  
A:Reference number: A32819; MUID:90001187; PMID:2477055  
A:Accession: A32819



A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 'MO', 25-176 <HOK>  
 A:Cross-references: GB:M27901; NID:9530166; PIDN:AAA42345.1; PID:9530167  
 R:Devos, A.; De Clercq, N.; Vercaren, I.; Heyns, W.; Rombauts, W.; Peeters, B.  
 Gene 125, 159-167, 1993  
 A:Title: Structure of rat genes encoding androgen-regulated cystatin-related proteins (C  
 A:Reference number: J70615; MUID:93216117; PMID:8462870  
 A:Accession: J70615  
 A:Molecule type: DNA  
 A:Residues: 1-156, 'V', 158-176 <DEV>  
 A:Cross-references: GB:M13993  
 R:Rindrickx, J.; Hemschoote, K.; De Clercq, N.; Van Dijk, P.; Peeters, B.; Rombauts, W.  
 Mol. Endocrinol. 4, 657-667, 1990  
 A:Title: Tissue-specific expression and androgen regulation of different genes encoding  
 A:Reference number: A34552; MUID:91125374; PMID:2280780  
 A:Accession: A34552  
 A:Molecule type: mRNA  
 A:Residues: 1-156, 'V', 158-176 <WIN>  
 A:Cross-references: GB:M58167; NID:9205902; PIDN:AAA63498.1; PID:9205903  
 R:Vercaren, I.; Winderickx, J.; Devos, A.; Peeters, B.; Heyns, W.  
 Endocrinology 132, 2496-2502, 1993  
 A:Title: An effect of androgens on the length of the poly(A)-tail and alternative splicing  
 A:Reference number: A49200; MUID:93178358; PMID:7679983  
 A:Note: Corrected and republished from Endocrinology 131, 2496-2502, 1992  
 A:Accession: A49200  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 137-156, 'V', 158-176 <VER>  
 A:Cross-references: EMBL:Z13993  
 A:Experimental source: Mistar albino  
 A:Note: sequence extracted from NCBI backbone (NCB1N.119368, NCB1P.119370)  
 C:Genetics:  
 A:Gene: Crp1  
 A:Introns: 79/3; 111/3; 149/3  
 C:Keywords: glycoprotein

Query Match 1.7%; Score 8; DB 2; Length 176;  
 Best Local Similarity 100.0%; Pred. No. 9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 421 TLULLAIF 428  
 |||||||  
 Db 8 TLULLAIF 15

RESULT 20  
 JC7335  
 chitinase (EC 3.2.1.14) 1 - cone shell (Conus tulipa)  
 C:Species: Conus tulipa (tulip cone)  
 C:Date: 08-Sep-2000 #sequence\_revision 08-Sep-2000 #text\_change 15-Sep-2000  
 C:Accession: JC7335  
 R:Yamagami, T.; Tsutsun, K.; Ishiguro, M.  
 Biosci. Biotechnol. Biochem. 64, 1394-1401, 2000  
 A:Title: Cloning, sequencing, and expression of the tulip bulb chitinase-1 cDNA.  
 A:Reference number: JC7335  
 A:Accession: JC7335  
 A:Molecule type: mRNA  
 A:Residues: 1-314 <YAM>  
 A:Cross-references: DDBJ:AB035668  
 C:Comment: This enzyme, a member of class III plant chitinases, which catalyzes the hydrolysis or protection against fungal pathogens.  
 C:Genetics:  
 A:Gene: tbc-1  
 C:Superfamily: alcohol sulfotransferase  
 C:Keywords: glycosidase; hydrolase

Query Match 1.7%; Score 8; DB 2; Length 314;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 169 SSULLLEF 176  
 |||||||

Db 7 SSULLLEF 14

RESULT 21  
 AF0205  
 Probable exported protein YP01686 [Imported] - Yersinia pestis (strain CO92)  
 C:Species: Yersinia pestis  
 C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001  
 C:Accession: AF0205  
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tlball, R.W.; Holden, M.T.G.; Prentice, M.  
 demo-Tarrega, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.  
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barril  
 Nature 413, 523-527, 2001  
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
 A:Reference number: AB0001; MUID:21470413; PMID:11586360  
 A:Accession: AF0205  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-317 <KUR>  
 A:Cross-references: GB:AL590842; PIDN:CAC90506.1; PID:915979716; GSPDB:GN00175  
 C:Genetics:  
 A:Gene: YP01686

Query Match 1.7%; Score 8; DB 2; Length 317;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 166 LISSLL 173  
 |||||||  
 Db 9 LISSLL 16

RESULT 22  
 H75253  
 hypothetical protein - Deinococcus radiodurans (strain R1)  
 C:Species: Deinococcus radiodurans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
 C:Accession: H75253  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.  
 M.; Shen, M.; Yamathavan, J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.;  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567266  
 A:Accession: H75253  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-347 <WHI>  
 A:Cross-references: GB:AE002089; GB:AE000513; NID:96460427; PIDN:AAF12144.1; PID:9646  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR2600  
 A:Map position: 1

Query Match 1.7%; Score 8; DB 2; Length 347;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 352 RSTPESRA 359  
 |||||||  
 Db 274 RSTPESRA 281

RESULT 23  
 T39599  
 conserved hypothetical protein SPCL65.07c - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 18-Feb-2000  
 C:Accession: T39599  
 R:Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.  
 Submitted to the EMBL Data Library, May 1998  
 A:Reference number: Z21866  
 A:Accession: T39599

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-354 <LYN>  
A:Cross-references: EMBL:AL023554; PIDN:CAA19027.1; GSPDB:GN00067; SPDB:SPBC16G5.07C  
A:Experimental source: strain 97zh; cosmid c16G5  
C:Genetics:  
A:Gene: SPDB:SPBC16G5.07C  
A:Map position: 2  
A:Insertions: 72/2; 265/2  
C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match  
Best Local Similarity 1.7%; Score 8; DB 2; Length 354;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 395 KASATASG 402  
|||||||  
DB 265 KASATASG 272

RESULT 24  
A84564  
hypothetical protein At2g18410 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: A84564  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.L.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Niemman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-392 <STO>  
A:Cross-references: GB:AE002093; NID:g4309726; PIDN:AAD15496.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g18410  
A:Map position: 2

Query Match  
Best Local Similarity 1.7%; Score 8; DB 2; Length 392;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 364 LSSSILAG 371  
|||||||  
DB 44 LSSSILAG 51

RESULT 25  
A97119  
probable membrane protein [imported] - Clostridium acetobutylicum  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C:Accession: A97119  
R:Rolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J.; Daly, M.T.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: A97119  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-581 <KUR>  
A:Cross-references: GB:AE001437; PIDN:AAK79740.1; PID:g15024745; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC1775

Query Match  
Best Local Similarity 1.7%; Score 8; DB 2; Length 581;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 KTYNAVAD 194  
|||||||  
DB 232 KTYNAVAD 239

RESULT 26  
A42322  
ornithine decarboxylase (EC 4.1.1.17) - Leishmania donovani  
C:Species: Leishmania donovani  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 24-Apr-1998  
C:Accession: A42322; S27821; S29187  
R:Hanson, S.; Adelman, J.; Ullman, B.  
J. Biol. Chem. 267, 2350-2359, 1992  
A:Title: Amplification and molecular cloning of the ornithine decarboxylase gene of L  
A:Reference number: A42322; MUID:92129311; PMID:1339439  
A:Accession: A42322  
A:Molecule type: DNA  
A:Residues: 1-707 <HAN>  
A:Cross-references: EMBL:M81192; NID:g159387; PID:g159388  
A:Note: sequence extracted from NCBI backbone (NCBIN:78350, NCBI:P:78355)  
C:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match  
Best Local Similarity 1.7%; Score 8; DB 2; Length 707;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 EVFKTYNV 191  
|||||||  
DB 168 EVFKTYNV 175

RESULT 27  
I53078  
homeotic gene regulator - mouse (fragment)  
C:Species: Mus sp. (mouse)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 02-Aug-2002  
C:Accession: I53078  
R:Randazzo, F.M.; Khavari, P.; Crabtree, G.; Tamkun, J.; Rosant, J.  
Dev. Biol. 161, 229-242, 1994  
A:Title: brg1: a putative murine homologue of the Drosophila brachna gene, a homeotic  
A:Reference number: I53078; MUID:94123856; PMID:7904967  
A:Accession: I53078  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1022 <RES>  
A:Cross-references: GB:S68108; NID:g545017; PIDN:ACG60670.1; PID:g545018  
C:Genetics:  
A:Gene: brg1  
A:Superfamily: human SNF2alpha protein; bromodomain homology  
F:860-915/Domain: bromodomain homology <BRO>

Query Match  
Best Local Similarity 1.7%; Score 8; DB 2; Length 1022;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 EODEEDE 71  
|||||||  
DB 663 EODEEDE 670

RESULT 28  
DVMS1A  
multidrug resistance protein 1a - mouse (fragment)  
N:Alternate names: P-glycoprotein 1a  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 19-Jan-2001  
C:Accession: A34175  
R:Hsu, S.I.H.; Lothstein, L.; Horwitz, S.B.  
J. Biol. Chem. 264, 12053-12062, 1989  
A:Title: Differential overexpression of three mdr gene family members in multidrug-re  
A:Reference number: A34175; MUID:89308614; PMID:2473069  
A:Accession: A34175

A:Molecule type: mRNA  
 A:Residues: 1-1104 <HSU>  
 A:Cross-references: GB:M24417; GB:J04839; NID:9200329; PIDN:AAA03243.1; PID:9200330  
 C:Comment: This is an integral membrane protein overproduced in multidrug-resistant cell  
 structurally and functionally unrelated lipophilic antitumor drugs.  
 C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
 C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; transmembrane pr  
 F:234-428/Domain: ATP-binding cassette homology <ABC1>  
 F:251-258/Region: nucleotide-binding motif A (P-loop)  
 F:375-379/Region: nucleotide-binding motif B  
 F:877-1073/Domain: ATP-binding cassette homology <ABC2>  
 F:994-901/Region: nucleotide-binding motif A (P-loop)  
 F:1020-1024/Region: nucleotide-binding motif B  
 F:1072/Binding site: ATP (Lys) #status predicted  
 F:900/Binding site: ATP (Lys) #status predicted

Query Match 1.7%; Score 8; DB 1; Length 1104;  
 Best Local Similarity 100.0%; Pred. No. 46;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 420 LTTLLAI 427  
 |||||  
 Db 681 LTTLLAI 688

## RESULT 29

## DWHYIC

multidrug resistance protein 1 - Chinese hamster

N:Alternate names: P-glycoprotein pgp1

C:Species: Cricetus griseus (Chinese hamster)

C>Date: 31-Dec-1990 #sequence\_revision 30-Sep-1992 #text\_change 19-Jan-2001

C:Accession: A38696; C38696; A27126; S33768; I52823

R:Devault, S.E.; Hussain, A.; Davide, J.P.; Melera, P.W.

J. Biol. Chem. 265: 4545-4555, 1991

A:Title: Full length and alternatively spliced pgp-1 transcripts in multidrug-resistant

A:Reference number: A38696; MOID:91154265; PMID:1571863

A:Accession: A38696

A:Molecule type: mRNA

A:Residues: 1-1276 <DEV>

A:Cross-references: GB:M59253; NID:9191154; PIDN:AAA37004.1; PID:9191155

A:Accession: C38696

A:Molecule type: mRNA

A:Residues: 108-1276 <DE1>

A:Cross-references: GB:M59254; NID:9191156; PIDN:AAA37005.1; PID:9191157

A:Experimental source: clone ADX185

A:Accession: B38696

A:Molecule type: mRNA

A:Residues: 1-32,771-1276 <DE2>

A:Cross-references: GB:M59252; NID:9191152; PIDN:AAA37003.1; PID:9191153

A:Experimental source: clone ADX124

R:Endicott, J.A.; Juranka, P.F.; Sarangi, F.; Gerlach, J.H.; Deuchars, K.L.; Ling, V.

Mol. Cell. Biol. 7, 4075-4081, 1987

A:Title: Simultaneous expression of two P-glycoprotein genes in drug-sensitive Chinese h

A:Reference number: A27126; MOID:88122132; PMID:2893255

A:Accession: A27126

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 706-1276 <END>

A:Cross-references: GB:M18797; NID:9191158; PIDN:AAA37006.1; PID:9191159

R:Zastawny, R.L.; Ling, V.

Biochim. Biophys. Acta 1173, 303-313, 1993

A:Title: Structural and functional analysis of 5' flanking and intron 1 sequences of the

A:Reference number: S33768; MOID:93305724; PMID:8100449

A:Accession: S33768

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-21 <ZAS>

A:Cross-references: EMBL:L03286

R:Teeter, L.D.; Eckersberg, T.; Tsai, Y.; Kuo, M.T.

Cell Growth Differ. 2, 426-437, 1991

A:Title: Analysis of the Chinese hamster P-glycoprotein/multidrug resistance gene pgp1

A:Reference number: I52823; MOID:92088970; PMID:1561134

A:Accession: I52823

A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-21 <RES>  
 A:Cross-references: GB:S81975; NID:9240862  
 C:Comment: This is an integral membrane protein overproduced in multidrug-resistant c  
 structurally and functionally unrelated lipophilic antitumor drugs.  
 C:Genetics: pgp1  
 C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
 C:Keywords: alternative splicing; ATP; duplication; glycoprotein; nucleotide binding;  
 F:407-601/Domain: ATP-binding cassette homology <ABC1>  
 F:424-431/Region: nucleotide-binding motif A (P-loop)  
 F:548-552/Region: nucleotide-binding motif B  
 F:1050-1246/Domain: ATP-binding cassette homology <ABC2>  
 F:1067-1074/Region: nucleotide-binding motif A (P-loop)  
 F:1193-1197/Region: nucleotide-binding motif B  
 F:87,91,96/Binding site: carbonylate (asn) (covalent) #status predicted  
 F:430/Binding site: ATP (Lys) #status predicted  
 F:1073/Binding site: ATP (Lys) #status predicted

Query Match 1.7%; Score 8; DB 1; Length 1276;  
 Best Local Similarity 100.0%; Pred. No. 53;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 420 LTTLLAI 427  
 |||||  
 Db 854 LTTLLAI 861

## RESULT 30

multidrug resistance protein 1a - mouse

C:Species: Mus musculus (house mouse)

C>Date: 13-Jul-1990 #sequence\_revision 13-Jul-1990 #text\_change 02-Feb-2001

C:Accession: A34786; A35671

R:Devault, A.; Gros, P.

Mol. Cell. Biol. 10, 1652-1663, 1990

A:Title: Two members of the mouse mdr gene family confer multidrug resistance with ov

A:Reference number: A34786; MOID:90205845; PMID:1969610

A:Accession: A34786

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1276 <DEV>

A:Cross-references: GB:M30697; NID:9199111; PIDN:AAA39517.1; PID:9387429

A:Experimental source: mouse mdra (P-glycoprotein) promoter reveals the

A:Reference number: A35671; MOID:90287150; PMID:1972547

A:Accession: A35671

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-525, 'HV', 528-938, 'S', 940-1035, 'V', 1037-1276 <HSU>

A:Cross-references: GB:M33581; NID:919104; PIDN:AAA39514.1; PID:9387427

C:Superfamily: multidrug resistance protein; ATP-binding cassette homology

C:Keywords: ATP; duplication; glycoprotein; membrane protein; nucleotide binding; P-1

F:406-600/Domain: ATP-binding cassette homology <ABC1>

F:423-431/Region: nucleotide-binding motif A (P-loop)

F:547-551/Region: nucleotide-binding motif B

F:1049-1245/Domain: ATP-binding cassette homology <ABC2>

F:1066-1074/Region: nucleotide-binding motif A (P-loop)

F:1192-1196/Region: nucleotide-binding motif B

F:429/Binding site: ATP (Lys) #status predicted

F:1072/Binding site: ATP (Lys) #status predicted

Query Match 1.7%; Score 8; DB 2; Length 1276;  
 Best Local Similarity 100.0%; Pred. No. 53;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 420 LTTLLAI 427  
 |||||  
 Db 853 LTTLLAI 860

```

RESULT 31
|DVH01
|multidrug resistance protein 1 - human
|N:Alternate names: P-glycoprotein 1
|C:Species: Homo sapiens (man)
|C>Date: 31-Dec-1990 #sequence_revision 18-Aug-1995 #text_change 19-Jan-2001
|C:Accession: A34914; PS0162; S15500; A25059; S43838; I52238; I65204
|R:Chen, C.; Clark, D.; Ueda, K.; Pastan, I.; Gottesman, M.M.; Roninson, I.B.
|J. Biol. Chem. 265, 506-514, 1990
|A:Title: Genomic organization of the human multidrug resistance (MDR1) gene and origin
|A:Reference number: A34914; M0ID:90094448; PMID:1967175
|A:Accession: A34914
|A:Molecule type: DNA
|A:Residues: 1-1280 <CH2>
|A:Cross-references: GB:M29447; GB:J05168; NID:g187496; PIDN:AAA59576.1; PID:g386862
|R:Kioka, N.; Yamano, Y.; Komano, T.; Ueda, K.
|submitted to JIPID, April 1991
|A:Reference number: PS0162
|A:Accession: PS0162
|A:Molecule type: DNA
|A:Residues: 1-22 <K10>
|R:Kioka, N.; Yamano, Y.; Komano, T.; Ueda, K.
|submitted to the EMBL Data Library, April 1991
|A:Description: transcriptional regulation of multidrug resistance gene (MDR1) expression
|A:Reference number: S15500
|A:Accession: S15500
|A:Molecule type: DNA
|A:Residues: 1-22, 'R' <K12>
|A:Cross-references: EMBL:X58723; NID:g34522; PIDN:CA41558.1; PID:g34523
|R:Chen, C.; Chin, J.E.; Ueda, K.; Clark, D.P.; Pastan, I.; Gottesman, M.M.; Roninson, I.
|Cell 47, 381-389, 1986
|A:Title: Internal duplication and homology with bacterial transport proteins in the mdr1
|A:Reference number: A25059; M0ID:87028230; PMID:2876781
|A:Accession: A25059
|A:Molecule type: mRNA
|A:Residues: 1-184, 'V', 186-1280 <CH2>
|A:Cross-references: GB:M14758; NID:g187468; PIDN:AAA59575.1; PID:g307180
|R:Chambers, T.C.; Pohl, J.; Glass, D.B.; Kuo, J.F.
|Biochem. J. 299, 309-315, 1994
|A:Title: Phosphorylation by protein kinase C and cyclic AMP-dependent protein kinase of
|A:Reference number: S43838; M0ID:94220047; PMID:7909431
|A:Accession: S43838
|A:Molecule type: protein
|A:Residues: 656-689 <CHA>
|R:Geckeler, V.; Weger, S.; Probst, H.
|Biochem. Biophys. Res. Commun. 169, 796-802, 1990
|A:Title: mdr1/P-glycoprotein gene segments analyzed from various human leukemic cell lin
|A:Reference number: I52238; M0ID:90290529; PMID:1972623
|A:Accession: I52238
|A>Status: translated from GB/EMBL/DBJ
|A:Molecule type: DNA
|A:Residues: 178-215 <RES>
|A:Cross-references: GB:M37724; NID:g183537; PIDN:AAA8047.1; PID:g553314
|A:Accession: I65204
|A>Status: translated from GB/EMBL/DBJ
|A:Molecule type: DNA
|A:Residues: 800-856 <RE2>
|A:Cross-references: GB:M37725; NID:g183538; PIDN:AAA8048.1; PID:g553315
|C:Comment: This is an integral membrane protein overproduced in multidrug-resistant cell
|structurally and functionally unrelated lipophilic antitumor drugs.
|C:Genetics:
|A:Gene: GDB:PGY1; MDRI
|A:Cross-references: GDB:120712; OMIM:171050
|A:Map position: 7q21-7q21
|C:Superfamily: multidrug resistance protein: ATP-binding cassette homology
|C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; phosphoprotein;
|F:1-658,553-1280/Region: duplication
|F:49-350/Domain: hydrophobic <HB1>
|F:351-637/Domain: hydrophilic <HL1>
|F:410-604/Domain: ATP-binding cassette homology <ABC1>
|F:427-434/Region: nucleotide-binding motif A (P-loop)
|F:551-555/Region: nucleotide-binding motif B
|F:638-708/Domain: linker <LIN>

```

```

F:709-993/Domain: hydrophobic <HB2>
F:994-1280/Domain: hydrophilic <HL2>
F:1053-1249/Domain: ATP-binding cassette homology <ABC2>
F:1070-1077/Region: nucleotide-binding motif A (P-loop)
F:1196-1200/Region: nucleotide-binding motif B
F:91,94,99/Binding site: carbonylate (Asn) (covalent) #status predicted
F:433/Binding site: ATP (Lys) #status predicted
F:661,667,671/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status
F:667,671,683/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #st
F:1076/Binding site: ATP (Lys) #status predicted

```

## Query Match

```

Best Local Similarity 1.7%; Score 8; DB 1; Length 1280;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 420 LTTLLAI 427
DB 857 LTTLLAI 864

```

## RESULT 32

```

S39059
protein Brg1 - human
C:Species: Homo sapiens (man)
C>Date: 25-Feb-1994 #sequence_revision 17-Nov-1995 #text_change 02-Aug-2002
C:Accession: S39059
R:Khavari, P.A.; Peterson, C.L.; Tankun, J.W.; Mendel, D.B.; Crabtree, G.R.
Nature 366, 170-174, 1993
A:Title: Brg1 contains a conserved domain of the SM12/SNF2 family necessary for norma
A:Reference number: S39059; M0ID:94050144; PMID:8232556
A:Accession: S39059
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1613 <KHA>
C:Superfamily: human SNF2alpha protein; bromodomain homology
F:1451-1506/Domain: bromodomain homology <BRO>

```

## Query Match

```

Best Local Similarity 1.7%; Score 8; DB 2; Length 1613;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 64 EDPDEDE 71
DB 1253 EDPDEDE 1260

```

## RESULT 33

```

A39401
merozoite surface antigen 1 precursor - Plasmodium vivax
C:Species: Plasmodium vivax
C>Date: 28-Feb-1992 #sequence_revision 28-Feb-1992 #text_change 24-Nov-1999
C:Accession: A39401
R:del Portillo, H.A.; Longacre, S.; Khouri, E.; David, P.H.
Proc. Natl. Acad. Sci. U.S.A. 88, 4030-4034, 1991
A:Title: Primary structure of the merozoite surface antigen 1 of Plasmodium vivax rev
A:Reference number: A39401; M0ID:91219506; PMID:2023952
A:Accession: A39401
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1726 <DER>
A:Cross-references: GB:M60807
C:Superfamily: major merozoite surface antigen
C:Keywords: surface antigen

```

## Query Match

```

Best Local Similarity 1.7%; Score 8; DB 2; Length 1726;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 173 LFFFSFI 180
DB 4 LFFFSFI 11

```

RESULT 34  
A56388  
dopamine- and cAMP-regulated neuronal phosphoprotein DARPP-32 - rat (fragments)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 03-Nov-1995 #sequence\_revision 03-Nov-1995 #text\_change 17-Mar-1999  
C:Accession: A56388  
R:Desolits, F.; Cohen, D.; Malm, A.C.; Greengard, P.; Girault, J.A.  
J. Biol. Chem. 270, 8772-8778, 1995  
A:Title: Phosphorylation of DARPP-32, a dopamine- and cAMP-regulated phosphoprotein, by  
A:Reference number: A56388; MUID:95238371; PMID:7721783  
A:Accession: A56388  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-49 <DESS>  
C:Keywords: phosphoprotein  
F:7,35,43/Binding site: phosphate (Ser) (covalent) (by casein kinase I) #status expirime

Query Match 1.5%: Score 7; DB 2; Length 49;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 66 DEEDEE 72  
DB 26 DEEDEE 32

RESULT 35  
S59962  
antimicrobial peptide gaegurin 5 precursor - Korean frog  
C:Species: Rana rugosa (Korean frog)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 16-Jul-1999  
C:Accession: S59962; PC2304  
R:Park, J.M.; Lee, J.Y.; Moon, H.M.; Lee, B.J.  
Biochim. Biophys. Acta 1264, 23-25, 1995  
A:Title: Molecular cloning of cDNAs encoding precursors of frog skin antimicrobial pepti  
A:Reference number: S59961; MUID:96038814; PMID:7578251  
A:Accession: S59962  
A:Molecule type: mRNA  
A:Residues: 1-65 <PAR>  
A:Cross-references: EMBL:U02393; NID:9733137; PIDN:AA64412.1; PID:9733138  
A:Experimental source: skin  
R:Park, J.M.; Jung, J.E.; Lee, B.J.  
Biochem. Biophys. Res. Commun. 205, 948-954, 1994  
A:Title: Antimicrobial peptides from the skin of a Korean frog, Rana rugosa.  
A:Reference number: PC2300; MUID:95091844; PMID:7999137  
A:Accession: PC2304  
A:Molecule type: protein  
A:Residues: 42-57, 'K', 59-65 <PAW>  
A:Experimental source: skin  
C:Comment: This peptide has antimicrobial activity.  
C:Superfamily: ranalexin precursor; dermorphin precursor amino-terminal homology  
F:1-41/Domain: dermorphin precursor amino-terminal homology <DER>  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-41/Domain: propeptide #status predicted <PRO>  
F:44-65/Product: antimicrobial peptide gaegurin 5 #status experimental <MAT>  
F:59-65/Region: rana box motif

Query Match 1.5%: Score 7; DB 2; Length 65;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 170 SLTLFF 176  
DB 7 SLTLFF 13

RESULT 36  
A53744  
ranalexin precursor - bullfrog  
C:Species: Rana catesbeiana (bullfrog)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A53744

R.Clark, D.P.; Durell, S.; Maloy, W.L.; Zasloff, M.  
J. Biol. Chem. 269, 10849-10855, 1994  
A:Title: Ranalexin. A novel antimicrobial peptide from bullfrog (Rana catesbeiana) sk  
A:Reference number: A53744; MUID:94193792; PMID:8144672  
A:Accession: A53744  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: protein  
A:Residues: 1-66 <CLAB>  
A:Cross-references: GB:S69903; NID:9546211; PIDN:AB30394.1; PID:9546212  
C:Superfamily: ranalexin precursor; dermorphin precursor amino-terminal homology  
F:1-46/Domain: dermorphin precursor amino-terminal homology <DER>  
F:47-66/Product: ranalexin #status experimental <MAT>  
F:60-66/Disulfide bonds: #status experimental

Query Match 1.5%: Score 7; DB 1; Length 66;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 170 SLTLFF 176  
DB 7 SLTLFF 13

RESULT 37  
LMHMA3  
H+-transporting two-sector ATPase (EC 3.6.3.14) lipid-binding protein - thermophilic  
N:Alternate names: H+-transporting ATP synthase chain c; hydrogen ion-transporting At  
C:Species: Thermophilic bacterium PS-3  
C:Date: 31-Jul-1979 #sequence\_revision 23-Oct-1981 #text\_change 14-Dec-2001  
C:Accession: S02255; A01048  
R:Ohta, S.; Yoshida, M.; Ishizuka, M.; Hirata, H.; Hamamoto, T.; Otawara-Hamamoto, Y.;  
Biochim. Biophys. Acta 933, 141-155, 1988  
A:Title: Sequence and over-expression of subunits of adenosine triphosphate synthase  
A:Reference number: S01397; MUID:88163679; PMID:2894854  
A:Accession: S02255  
A:Molecule type: DNA  
A:Residues: 1-72 <OHR>  
A:Cross-references: EMBL:X07804; NID:945808; PIDN:CA30649.1; PID:945811  
A:Note: this sequence was confirmed by protein sequencing  
R:Hoppe, J.; Sebald, W.  
Eur. J. Biochem. 107, 57-65, 1980  
A:Title: Amino acid sequence of the proteolipid subunit of the proton-translocating A  
A:Reference number: A01048; MUID:80245993; PMID:6447066  
A:Accession: A01048  
A:Molecule type: protein  
A:Residues: 1-72 <HOP>  
C:Comment: This protein is a nonenzymatic component of ATPase.  
C:Superfamily: H+-transporting ATP synthase lipid-binding protein  
C:Keywords: ATP biosynthesis; hydrolase; membrane-associated complex; transmembrane p  
F:1-72/Product: H+-transporting ATP synthase lipid-binding protein #status experiment  
F:2-28/Domain: transmembrane #status predicted <TM1>  
F:46-72/Domain: transmembrane #status predicted <TM2>  
F:1/Modified site: N-formylmethionine #status experimental  
F:56/Active site: Glu #status predicted

Query Match 1.5%: Score 7; DB 1; Length 72;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 177 FSFIYLG 183  
DB 65 FSFIYLG 71

RESULT 38  
C71978  
hypothetical protein jhp0055 - Helicobacter pylori (strain J99)  
C:Species: Helicobacter pylori  
A:Variety: strain J99  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 08-Oct-1999  
C:Accession: C71978  
R:Alm, R.A.; Ling, L.S.L.; Molr, D.T.; King, B.L.; Brown, E.D.; Dolg, P.C.; Smith, D.

Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;  
 Nature 397, 176-180, 1999  
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
 A:Reference number: A71800; MID:9910557; PMID:9923682  
 A:Accession: C71978  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-78 <ARN>  
 A:Cross-references: GB:AE001445; GB:AE001439; NID:g4154559; PIDN:AMD05639.1; PID:g415456  
 A:Experimental source: strain J99  
 C:Genetics:  
 A:Gene: jnp0055

Query Match 1.5%; Score 7; DB 2; Length 78;  
 Best Local Similarity 100.0%; Pred. No. 41;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 109 KDGOLIV 115  
 |||||  
 Db 65 KDGOLIV 71

## RESULT 39

B47354  
 ygdF protein - Salmonella typhimurium (fragment)  
 C:Species: Salmonella typhimurium  
 C:Date: 08-May-1995 #sequence\_revision 12-May-1995 #text\_change 08-Oct-1999  
 C:Accession: B47354  
 R:Can, K.; Gupta, S.D.; Sankaran, K.; Schmid, M.B.; Wu, H.C.  
 J. Biol. Chem. 268, 16544-16550, 1993  
 A:Title: Isolation and characterization of a temperature-sensitive mutant of Salmonella  
 A:Reference number: A47354; MID:93346403; PMID:8344935  
 A:Accession: B47354  
 A:Status: preliminary; translation not shown; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-85 <GAN>  
 A:Cross-references: GB:LI3259; NID:g295201; PIDN:AAA20895.1; PID:g511503  
 C:Genetics:  
 A:Gene: ygdF  
 C:Superfamily: phosphotransferase system enzyme I homology (fragment) <PI1>  
 F:1-64/Domain: phosphotransferase system enzyme I homology (fragment) <PI1>

Query Match 1.5%; Score 7; DB 2; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 44;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 412 VAILIGL 418  
 |||||  
 Db 15 VAILIGL 21

## RESULT 40

C71549  
 hypothetical protein CT164 - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
 C:Species: Chlamydia trachomatis  
 C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999  
 C:Accession: C71549  
 R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,  
 Science 282, 754-759, 1998  
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra  
 A:Reference number: A71570; MID:9900809; PMID:9784136  
 A:Accession: C71549  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-86 <ARN>  
 A:Cross-references: GB:AE001290; GB:AE001273; NID:g3328560; PIDN:AAC67755.1; PID:g332856  
 A:Experimental source: serotype D, strain UW-3/Cx  
 C:Genetics:  
 A:Gene: CT164

Query Match 1.5%; Score 7; DB 2; Length 86;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 168 ISSULLL 174  
 |||||  
 Db 46 ISSULLL 52

## RESULT 41

AF0856  
 conserved hypothetical protein SRY3056 [imported] - Salmonella enterica subsp. enteri  
 C:Species: Salmonella enterica subsp. enterica serovar typh  
 A:Note: this species has also been called Salmonella typh  
 C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 27-Nov-2001  
 C:Accession: AF0856  
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church  
 th, T.; Connor, P.; Croxin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr  
 , S.; Mole, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,  
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se  
 A:Reference number: AB0502; PMID:11677608  
 A:Accession: AF0856  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-103 <PAR>  
 A:Cross-references: GB:AL513382; PIDN:CAD06037.1; PID:g16504004; GSPDB:GN00176  
 C:Genetics:  
 A:Gene: SRY3056  
 C:Superfamily: hypothetical protein HI0673

Query Match 1.5%; Score 7; DB 2; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 420 LTLILLA 426  
 |||||  
 Db 4 LTLILLA 10

## RESULT 42

C83236  
 hypothetical protein PA3284 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: C83236  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warner, P.; Hickey, M.J.;  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L  
 .; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
 A:Reference number: AB2950; MID:20437337; PMID:10984043  
 A:Accession: C83236  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-104 <STO>  
 A:Cross-references: GB:AE004750; GB:AE004091; NID:g9949400; PIDN:AAG06672.1; GSPDB:GN  
 C:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA3284

Query Match 1.5%; Score 7; DB 2; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 53;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 360 AVOELSS 366  
 |||||  
 Db 31 AVOELSS 37

## RESULT 43

H72575  
 hypothetical protein APE1888 - Aeropyrum pernix (strain K1)  
 C:Species: Aeropyrum pernix  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999

C:Accession: H72575  
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takat  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
DNA Res. 6: 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
A:Reference number: A72450; MUID:9310339; PMID:10382966  
A:Accession: H72575  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-107 <RAW>  
A:Cross-references: DDBJ:AP000062; NID:95105244; PIDN:BAA80893.1; PID:dl044679; PID:9510  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE1888

Query Match 1.5%; Score 7; DB 2; Length 107;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 364 LSSSILA 370  
|||||  
Db 2 LSSSILA 8

RESULT 44  
D72617  
hypothetical protein APE1397 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jun-2000  
C:Accession: D72617  
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takat  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
DNA Res. 6: 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
A:Reference number: A72450; MUID:9310339; PMID:10382966  
A:Accession: D72617  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-108 <RAW>  
A:Cross-references: DDBJ:AP000061; NID:95104821; PIDN:BAA80394.1; PID:dl044180; PID:9510  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE1397  
C:Superfamily: Aeropyrum pernix hypothetical protein APE1397

Query Match 1.5%; Score 7; DB 2; Length 108;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 364 LSSSILA 370  
|||||  
Db 70 LSSSILA 76

RESULT 45  
B72763  
hypothetical protein APE0095 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jun-2000  
C:Accession: B72763  
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takat  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
DNA Res. 6: 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
A:Reference number: A72450; MUID:9310339; PMID:10382966  
A:Accession: B72763  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-109 <RAW>  
A:Cross-references: DDBJ:AP000058; NID:95103388; PIDN:BAA79004.1; PID:dl042780; PID:9510  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE0095

C:Superfamily: Aeropyrum pernix hypothetical protein APE0095

Query Match 1.5%; Score 7; DB 2; Length 109;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 147 TILLVL 153  
|||||  
Db 43 TILLVL 49

RESULT 46  
S31130  
hypothetical protein F59B2.10 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 04-Mar-2000  
C:Accession: S31130  
R:Stilson, J.; Du, Z.; Thomas, K.; Wilson, R.; Hillier, L.; Staden, R.; Halloran, N.;  
awkins, T.; Alnough, R.; Waterston, R.  
submitted to the EMBL Data Library, November 1991  
A:Description: The C. elegans sequencing project: A beginning.  
A:Reference number: S31122  
A:Accession: S31130  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-117 <SQL>  
A:Cross-references: EMBL:211505; NID:96718; PID:96727  
C:Genetics:  
A:Inserts: 21/2; 74/1  
C:Superfamily: Caenorhabditis elegans hypothetical protein F59B2.10

Query Match 1.5%; Score 7; DB 2; Length 117;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 168 ISSLLLL 174  
|||||  
Db 8 ISSLLLL 14

RESULT 47  
S61635  
ribosomal protein L22.e.A, cytosolic - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein L2168; protein YLR061w; ribosomal protein YL31  
C:Species: Saccharomyces cerevisiae  
C:Date: 09-Mar-1996 #sequence\_revision 12-Apr-1996 #text\_change 19-Apr-2002  
C:Accession: S61635; S64889; S11582  
R:Urestarazu, L.A.  
submitted to the EMBL Data Library, December 1995  
A:Reference number: S61618  
A:Accession: S61635  
A:Molecule type: DNA  
A:Residues: 1-121 <URR>  
A:Cross-references: EMBL:X94607; NID:91181264; PIDN:CAA64308.1; PID:91181282  
R:Andre, B.; Urestarazu, L.A.  
submitted to the Protein Sequence Database, May 1996  
A:Reference number: S64872  
A:Accession: S64889  
A:Molecule type: DNA  
A:Residues: 1-121 <AND>  
A:Cross-references: EMBL:Z71233; NID:91360407; PIDN:CAA97592.1; PID:91370303; MIPS:YL  
R:Okada, E.; Higo, K.; Osawa, S.  
Biochemistry 21, 4545-4550, 1982  
A:Title: Isolation of seventeen proteins and amino-terminal amino acid sequences of e  
A:Reference number: S11575; MUID:83048950; PMID:6814480  
A:Accession: S11582  
A:Molecule type: protein  
A:Residues: 2-17, 'B', 19-24, 'B', 26-28, 'B', 30-38, 'B', 40-43, 'Z', 45-48, 'B', 50-51, 'B' <OTA  
C:Genetics:  
A:Gene: SGD:RPL22A  
A:Cross-references: SGD:S0004051  
A:Map position: 12R

A: Introns: 4/3  
A: Note: YLR061w  
C: Superfamily: rat ribosomal protein L22  
C: Keywords: cytosol; protein biosynthesis; ribosome

Query Match 1.5%; Score 7; DB 2; Length 121;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 66 DEEDEE 72  
|||||  
DB 115 DEEDEE 121

RESULT 48  
E97486  
Hypothetical protein AGR\_C\_1913 [imported] - Agrobacterium tumefaciens (strain C58, Cere

C: Species: Agrobacterium tumefaciens  
C: Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
C: Accession: E97486

R: Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
A.; Liu, F.; Hljam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A: Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A: Reference number: A97359; PMID:11743194

A: Accession: E97486

A: Status: preliminary

A: Molecule type: DNA

A: Residues: 1-125 <KUR>

A: Cross-references: GB:AE007869; PIDN:AAK86846.1; PID:q15156060; GSPDB:GN00169

C: Genetics:

A: Gene: AGR\_C\_1913

A: Map position: circular chromosome

Query Match 1.5%; Score 7; DB 2; Length 125;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ROVEED 66  
|||||  
DB 41 ROVEED 47

RESULT 49  
AE2704  
Two component response regulator Atu1038 [imported] - Agrobacterium tumefaciens (strain

C: Species: Agrobacterium tumefaciens  
C: Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
C: Accession: AE2704

R: Wood, D.W.; Secubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I

erage, G.; Gillet, W.; Grant, C.; Gunthner, D.; Kutayav, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001

A: Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, B.W.  
A: Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A: Reference number: AB2577; PMID:11743193

A: Accession: AE2704

A: Status: preliminary

A: Molecule type: DNA

A: Residues: 1-125 <KUR>

A: Cross-references: GB:AE008688; PIDN:AAI42051.1; PID:q17739429; GSPDB:GN00186

A: Experimental source: strain C58 (Dupont)

C: Genetics:

A: Gene: Atu1038

A: Map position: circular chromosome

Query Match 1.5%; Score 7; DB 2; Length 125;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ROVEED 66  
|||||

DB 41 ROVEED 47

RESULT 50

D87491

Hypothetical protein CC1953 [imported] - Caulobacter crescentus

C: Species: Caulobacter crescentus

C: Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C: Accession: D87491

R: Niernan, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,

B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A: Title: Complete Genome Sequence of Caulobacter crescentus.

A: Reference number: AB7249; NCID:21173698; PMID:11259647

A: Accession: D87491

A: Status: preliminary

A: Molecule type: DNA

A: Residues: 1-131 <STO>

A: Cross-references: GB:AE005673; NID:q13423412; PIDN:AAK23928.1; GSPDB:GN00148

C: Genetics:

A: Gene: CC1953

Query Match 1.5%; Score 7; DB 2; Length 131;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 137 AIMISV 143  
|||||  
DB 52 AIMISV 58

Search completed: February 6, 2003, 14:21:09  
Job time : 62.0975 secs